Universität Basel
Fakultät für Psychologie

Cognitive assessment in healthy and pathological aging

Inauguraldissertation zur Erlangung der Würde eines Doktors der Philosophie vorgelegt der Fakultät für Psychologie der Universität Basel von

Alessandra Thomann

aus Gaiserwald, SG

Basel, 2019

Universität
Basel

Fakultät für
Psychologie

Universität Basel

Fakultät für Psychologie

Genehmigt von der Fakultät für Psychologie auf Antrag von

Prof. Dr. phil. Andreas U. Monsch

Prof. Dr. phil. Alexander Grob

Datum des Doktoratsexamen: 18.06.2019

DekanIn der Fakultät für Psychologie

Universität Basel
Fakultät für Psychologie

Erklärung zur wissenschaftlichen Lauterkeit

Ich erkläre hiermit, dass die vorliegende Arbeit ohne die Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel selbstständig verfasst habe. Zu Hilfe genommene Quellen sind als solche gekennzeichnet. Die veröffentlichten oder zur Veröffentlichung in Zeitschriften eingereichten Manuskripte wurden in Zusammenarbeit mit den Koautoren erstellt und von keinem der Beteiligten an anderer Stelle publiziert, zur Publikation eingereicht, oder einer anderen Prüfungsbehörde als Qualifikationsarbeit vorgelegt. Es handelt sich dabei um folgende Manuskripte:

- **Thomann AE**, Goettel N, Monsch RJ, Berres M, Jahn T, Steiner LA, Monsch AU. *The Montreal Cognitive Assessment: Normative Data from a German-speaking Cohort and Comparison with International Normative Samples*. J Alzheimers Dis. 2018;64(2):643-655. **Published.**

- **Thomann AE**, Berres M, Goettel N, Steiner LA, Monsch AU. *Two separate cut-offs on the MoCA for patients with a neurocognitive disorder.* **Submitted.**

- Monsch RJ, Burckhardt AC, Berres M, **Thomann AE**, Ehrensperger MM, Steiner LA, Goettel N. *Development of a Novel Self-administered Cognitive Assessment Tool and Normative Data for Older Adults. J* Neurosurg Anesthesiol. 2019;31(2):218-226. **Published.**

Basel, 01. Mai 2019

Alessandra Thomann

Content

Danksagung

Ganz besonders danken möchte ich meinem Supervisor und Mentor Prof. Dr. phil. Andreas Monsch, welcher während meiner Dissertation stets eine grossartige Unterstützung war, mich gefördert und gefordert hat, immer eine offene Tür für mich hatte und mich auch besonders bei der Planung meiner beruflichen Zukunft bestens beraten und unterstützt hat. Andreas Monsch hat mir fachlich sehr viel Vertrauen entgegengebracht und mir ausserordentliche Chancen ermöglicht, meine Kenntnisse weiterzuentwickeln. Ein grosses Dankeschön dafür.

Bei Herrn Prof. Dr. phil. Alexander Grob bedanke ich mich herzlich für die kompetente und freundliche Co-Betreuung meiner Doktorarbeit.

Mein Dank geht auch an Dr. med. Nicolai Goettel, welcher mich während meiner Dissertation im Universitätsspital Basel stets unterstützt und mich bei der Planung meiner weiteren beruflichen Karriere beraten hat. Weiter bedanke ich mich bei allen Co-Autoren der verfassten Publikationen für ihre konstruktive Mitarbeit und für die stets kollegiale und interessante Zusammenarbeit im Publikationsprozess.

Ganz herzlichen Dank an alle Mitarbeiter der Memory Clinic Basel, welche den Arbeitsalltag immer freundlich gestaltet haben und viel zu einem familiären Klima beigetragen haben. Besonders nennen möchte ich meine Büronachberinnen Dr. phil. Sabine Krumm, Jael Fasnacht und Marlen Frei, mit welchen ich mich sowohl beruflich als auch privat austauschen konnte. Weiter geht mein Dank an Dr. phil. Stefan Bläsi, welcher mich beim Entscheid die Dissertation zu beginnen, bestens beraten hat.

Ein grosser Dank geht auch an meine Familie und Freunde, welche mich stets begleitet und unterstützt haben. Ganz besonders danken möchte ich an dieser Stelle meinem besten Freund und Lebenspartner Christian Brauchli, welcher mich in zahlreichen Diskussionen fachlich unterstützt hat, mich immer wieder angespornt hat, das Beste aus mir herauszuholen und mir vor allem auch persönlich und in schwierigeren Zeiten eine sehr wertvolle Stütze war.

Abstract

The current demographical development leads to a growing number of older people and thereby to an increase in patients suffering from age-related diseases like dementia. Facing this healthcare challenge requires an efficient and accurate diagnostic process. An understanding of healthy cognitive aging is essential to recognize and investigate pathological states. Further the diagnostic accuracy of a given tool needs to be assessed in the clinical setting where it is ultimately applied. The aim of the present doctoral thesis is to provide insights into the cognitive performance of cognitively healthy older individuals and to assess the diagnostic accuracy of a well-known tool in a realistic clinical routine setting.

We found demographic-related effects on cognitive performance in 283 cognitively healthy individuals who were assessed with two different cognitive assessment tools: the Montreal Cognitive Assessment (MoCA) and a newly developed computerized cognitive assessment (CogCheck). Adjusting for these effects by converting raw scores to standard scores, lead to higher specificity of the MoCA.

In a second study, we investigated the diagnostic accuracy of the original MoCA cut-off (25/26) in a clinical routine setting to differentiate cognitive normal findings from patients with a neurocognitive disorder (NCD; N = 496). While the original cut-off yielded high sensitivity, its specificity was poor. The classification rate increased when a lower cut-off score (23/24) was applied. However, sensitivity to detect mild NCD was low. We therefore proposed a new way to evaluate cognitive performance: Combining two separate cut-offs (23/24 and 26/27) with a gray area allows for both, high specificity and high sensitivity. Additional examinations are required in the gray area between these two cut-offs.

Finally, we have found important heterogeneities in the methodology of cognitive normative studies. This information may guide future endeavors to create guidelines for the definition of cognitive health, which is a baseline requirement to investigate pathological changes. As an outlook, methodological reflections on the evaluation of cognitive assessments are given and the role of neuropsychology in the age of digitalization is discussed.

1. General introduction

The current demographical development is characterized by a growing number of older individuals and thereby an increase in patients suffering from age-related diseases. For instance, dementia cases are estimated to nearly triple and reach 131.5 million patients by the year 2050 (Prince, Comas-Herrera, Knapp, Guerchet, & Karagiannidou, 2016). Dementia as a clinical syndrome is characterized by cognitive impairment that interferes with activities of daily living and represents a decline from a previous level of functioning. It may be caused by a variety of underlying etiologies with Alzheimer's disease (AD) being the leading one (Winblad et al., 2016). While some forms of dementia and cognitive impairment are potentially reversible if treated appropriately (Clarfield, 2003), there are currently no satisfying care options for neurodegenerative diseases like AD. Incipient pharmaceutical or non-pharmaceutical therapies target early stages of the disease (Scheltens et al., 2016), making early-detection of cognitive impairment crucial. Moreover, an early implementation of current treatment strategies may slow progression of cognitive decline, allows the treatment of secondary behavioral or psychiatric symptoms and the organization of care support; thereby increasing the patients' and their caregivers' quality of life (Petersen et al., 2017). However, diagnosing AD (and other neurodegenerative diseases) is still challenging and it is especially difficult in earlier disease stages when only subtle symptoms are apparent.

The growth of the geriatric patient population furthermore has an impact on the hospital setting. The need for surgical procedures increases with age (Hall, DeFrances, Williams, Golosinskiy, & Schwartzman, 2010) and older people have a higher risk for adverse postoperative cognitive outcomes like postoperative delirium (POD) or postoperative cognitive dysfunction (POCD) (Story et al., 2010). These cognitive disorders are in turn associated with higher morbidity and mortality (Sanders, Pandharipande, Davidson, Ma, & Maze, 2011; Steinmetz, Christensen, Lund, Lohse, & Rasmussen, 2009; Witlox et al., 2010). Individuals with a higher risk for POD or POCD may benefit from preventive measures or increased postoperative care (Inouye et al., 1999), making preoperative identification of risk factors essential.

Consequently, the increasing number of geriatric patients is associated with important health-related, economic, and social challenges. Addressing these

confronts is a priority in healthcare. The field of neuropsychology, which investigates brain-behavior relationships, offers important contributions to this aim. Cognitive dysfunction is the primary deficit in dementia (*American Psychiatric Association*, 2013) and cognitive assessment plays an essential role for differential diagnosis, for the assessment of disease severity, for predictions on the disease course, and as a measure of treatment success. In the context of adverse postoperative outcomes, pre-existing cognitive impairment is one of the leading risk factors (Dasgupta & Dumbrell, 2006; Inouye, Westendorp, & Saczynski, 2014; Jones et al., 2016; Nadelson, Sanders, & Avidan, 2014; Silbert et al., 2015; Sprung et al., 2017). Preoperative cognitive assessment may therefore help identifying high-risk surgery patients who could benefit from increased pre- and post-operative care.

1.1 Cognitive assessment: contributions and challenges

In the 5th version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; *American Psychiatric Association,* 2013), disorders with acquired cognitive impairment as the leading clinical symptom are referred to as neurocognitive disorder (NCD). NCDs are further divided in two levels of severity, with the previously introduced term dementia being referred to as major NCD. Major NCD is characterized by a decline of more than two standard deviations (SD) from a healthy normative population in at least one cognitive domain. Further, the cognitive deficits interfere with independent functioning in everyday life. A minor form of NCD is entitled mild NCD and relates to the concept of Mild cognitive impairment (MCI; Petersen, 2004). Patients with mild NCD typically score in the range of one to two standard deviations (SD) from a healthy normative population in at least one cognitive domain and are still independent in everyday activities. Of note, there is no clear border between these two entities since cognitive decline represents a continuum rather than distinct categories. NCD may be caused by a variety of underlying diseases (i.e., AD, frontotemporal lobar degeneration, Lewy body disease, vascular disease, traumatic brain injury, substance/medication use, human immunodeficiency virus infection, prion disease, Parkinson's disease, Huntington's disease, other medical conditions, or multiple etiologies). The diagnostic criteria for mild and major NCD are based on the following six cognitive domains that should be evaluated in a comprehensive neuropsychological

assessment: Perceptual-motor function, executive function, complex attention, language, social cognition, learning and memory. A comprehensive assessment should include at least two different measures per domain.

Many currently used cognitive assessment tools have been developed decades ago (e.g., Mini-Mental State Examination in 1975 [Folstein, Folstein, & McHugh, 1975], Consortium to Establish a Registry for Alzheimer's Disease – Neuropsychological Assessment Battery in 1989 [Morris et al., 1989], Stroop-Test in 1935 [Stroop, 1935]). Therefore, new evidence on brain-behavior-relationships is often not considered in standard cognitive assessments. This may partly be due to the lack of available normative data for more recent cognitive assessment tools. The evaluation of cognition requires a definition of what is normal and a concept of deviations that constitute an impairment. Additionally, the performance in cognitive tests may be influenced by demographic characteristics like age, education, and sex (Casaletto & Heaton, 2017). Further, culture and language may have an impact on item-difficulty. Therefore, it is important to investigate the performance of healthy individuals that are comparable to the target population, before applying a cognitive test in a clinical setting. Conducting large-scale normative studies for every new test and for many specific populations would be ideal. However, this type of studies is expensive and is usually not financially supported (Casaletto & Heaton, 2017). Consequently, representable norms for the patient population are often lacking and/or the available norms have been developed based on previous generations of individuals, which may be outdated.

Like every diagnostic test, cognitive assessment tools should have sufficient validity and high diagnostic accuracy. False-negative diagnoses due to a lack of sensitivity deprives patients from access to treatment or clinical trials. On the other hand, false-positive diagnoses due to poor specificity leads to avoidable stress and burden for a patient, costs due to unnecessary examinations and treatments, and inappropriate inclusion in clinical trials. Since in the context of neurodegenerative diseases there is no clear benefit of favoring false-positives over false-negatives or vice-versa, sensitivity and specificity should be balanced.

1.2 Diagnostic steps for cognitive impairment and dementia

In the diagnostic workup for dementia, patients are usually first seen by a general practitioner (GP). At this level, a case-finding approach rather than broad screening has been recommended (Ehrensperger et al., 2014), meaning that brief cognitive assessment tools should only be applied in those individuals that present with red flags indicative of possible cognitive impairment (e.g., report of cognitive worsening by the patient or an informant; signs of cognitive worsening that become apparent to the clinician during routine examination). A pathological result in a first-step test is usually followed by a referral to a specialized clinic where extensive neuropsychological and medical examinations take place. A comprehensive dementia workup should include detailed patient and medical history—if possible combined with reports from an informant—, a comprehensive neuropsychological assessment, a neurological and geriatric evaluation, laboratory diagnostics, brain imaging (magnetic resonance imaging [MRI]; positron emission tomography [PET]), and sometimes the assessment of protein depositions in cerebrospinal fluid or PET (Frisoni et al., 2017).

In conclusion, the diagnostic process of dementia is a multi-disciplinary workup, it includes identification processes at different levels (i.e. the GP level and the specialist level), it is time-consuming, expensive, personnel-intensive, and sometimes invasive (e.g. lumbar puncture to collect CSF). Additionally, the number of patients with dementia are rapidly increasing and specialized clinics are already facing long waiting lists. Thus, it is crucial to apply tools with high diagnostic accuracy to detect those patients that should benefit from such extensive assessments while at the same time filter out healthy individuals that should not undergo unnecessary examinations. Ideally, such first-step tools should be brief and inexpensive, and their administration should not require highly trained personnel. This becomes even more important, if there should one day be a treatment for AD with a significant positive effect. It is reasonable to imagine, that in this scenario, waiting lists would drastically increase.

Therefore, one main challenge of the field is finding ways to improve the efficiency of the diagnostic process while still providing high diagnostic accuracy. This may be achieved by targeting (a) current screening procedures at the GPs office, and (b) examinations at the specialized level. A more efficient assessment of pre-

existing cognitive impairment may also be beneficial in the pre-surgery setting where time is limited and trained neuropsychologists are usually absent.

1.2.1 First level: General practitioners' office

On the GP level, many screening tools exist to briefly assess for cognitive impairment. In this context, the Mini-Mental state examination (MMSE) is probably the most-known test and it has been used during decades for cognitive screening. However, like many other cognitive tools, the MMSE has been developed in the 70ies and only poorly assesses the six cognitive domains proposed in the DSM-5 (2013). There is growing evidence, that the MMSE has poor sensitivity, especially to detect subtle cognitive deficits that are present in MCI (Ciesielska et al., 2016; Nasreddine et al., 2005; Roalf et al., 2013). To address this issue, other screening tools have been developed, some of them with a special focus on the detection of MCI like the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005). Compared to the MMSE, the MoCA has superior diagnostic accuracy for MCI (Ozer, Young, Champ, & Burke, 2016; Trzepacz et al., 2015), correlates better with extensive neuropsychological test batteries (Lam et al., 2013), and covers most of the cognitive domains outlined in the DSM-5 (2013). Namely, executive functions and complex attention are assessed, which are not considered in the MMSE. While the MoCA gains popularity as an alternative to the MMSE, there are still some concerns that need to be addressed. The initially proposed MoCA cut-off score (25/26 points) has poor specificity (Carson, Leach, & Murphy, 2018; Davis et al., 2015) and demographical characteristics are not appropriately considered. So far, no study has investigated the properties of the German MoCA in cognitively healthy individuals and patients, why most German-speaking clinicians still rely on this cut-off. With false-positive rates ranging from 46% (Malek-Ahmadi et al., 2015) up to 76% (Rossetti et al., 2017) in other MoCA normative studies, the use of this cut-off may seriously decrease the efficiency of the case-finding process and may lead to many false-positive referrals to specialized clinics. Therefore, in study I and II we aimed at investigating the properties of the German MoCA and deducting ways to decrease the false-positive rate while at the same time keeping sensitivity for cognitive impairment high. In study I, we assessed the MoCA performance in cognitively healthy individuals and investigated whether age, education, and/or sex have an impact on MoCA scores. In study II, we completed this knowledge with patient data and analyzed ways to increase the diagnostic accuracy of the MoCA to distinguish healthy individuals form patients with mild or major NCD.

1.2.2 Second level: Comprehensive cognitive assessment

In earlier years of neuropsychological assessment, cognitive functions were usually evaluated in a qualitative way, which tailors the examination to the needs and characteristics of a specific patient (Casaletto & Heaton, 2017). Over the years, quantitative assessments have gained popularity to increase comparability and reproducible results. Today, fixed batteries that include tests on the most important cognitive domains are often performed as a standard in all patients. A standard assessment may then be combined with more specific tests and in-depth assessments if needed. Generally, a comprehensive neuropsychological assessment is time-consuming and takes up to two hours for the test administration plus additional time for test scoring and interpretation. Additionally, trained personnel is required to perform and interpret the assessments.

Considering the availability of modern technologies, computerized cognitive assessment tools gain increasing attention and are a potential way to increase the efficiency of this process. Especially in a setting, where many individuals should be assessed for cognitive impairment, a computerized cognitive assessment that may be performed without the assistance of a trained professional yields interesting possibilities and may reduce costs.

However, before a computerized assessment tool can be applied in a clinical setting, it should undergo the same development and validation steps as traditional paper-and-pencil-tests. It needs to be assessed for feasibility in the target population, normative data should be developed, and its diagnostic accuracy must be investigated. In a joint-collaboration between the Department of Anesthesia of the University Hospital Basel and the Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel, a new self-administered computerized cognitive assessment tool (CogCheck) has been developed and tested for feasibility in two pilot-studies (Anyiam, 2018; Burckhardt, 2014). In study III, CogCheck was administered to cognitively healthy individuals to assess the effect of age, education, and sex on the CogCheck performance and to provide normative values for the tool.

2. List of publications

I. **Thomann AE**, Goettel N, Monsch RJ, Berres M, Jahn T, Steiner LA, Monsch AU. *The Montreal Cognitive Assessment: Normative Data from a German-speaking Cohort and Comparison with International Normative Samples*. J Alzheimers Dis. 2018;64(2):643-655. **Published.**

II. **Thomann AE**, Berres M, Goettel N, Steiner LA, Monsch AU. *Two separate cut-offs on the MoCA for patients with a neurocognitive disorder*. **Submitted.**

III. Monsch RJ, Burckhardt AC, Berres M, **Thomann AE**, Ehrensperger MM, Steiner LA, Goettel N. *Development of a Novel Self-administered Cognitive Assessment Tool and Normative Data for Older Adults. J* Neurosurg Anesthesiol. 2019;31(2):218-226. **Published.**

3. Study I:

The Montreal Cognitive Assessment: Normative Data from a German-speaking Cohort and Comparison with International Normative Samples

Alessandra E. Thomann[a,b,1], Nicolai Goettel[b,c,1], Raphael J. Monsch[b], Manfred Berres[d], Thomas Jahn[e], Luzius A. Steiner[b,c], Andreas U. Monsch[a]

Published

[a]Memory Clinic, University Center for Medicine of Aging, Felix Platter Hospital, Basel, Switzerland
[b]Department of Anesthesia, Surgical Intensive Care, Prehospital Emergency Medicine and Pain Therapy, University Hospital Basel, Basel, Switzerland
[c]Department of Clinical Research, University of Basel, Basel, Switzerland
[d]Department of Mathematics and Technology, University of Applied Sciences Koblenz, Koblenz, Germany
[e]Department of Psychiatry and Psychotherapy, Technische Universität München, Klinikum rechts der Isar, Munich, Germany

[1]Alessandra E. Thomann and Nicolai Goettel contributed equally to the work and are co-first authors of this paper.

The final publication is available at IOS Press through
http://dx.doi.org/10.3233/JAD-180080

ABSTRACT

Background: The Montreal Cognitive Assessment (MoCA) is used to evaluate multiple cognitive domains in elderly individuals. However, it is influenced by demographic characteristics that have yet to be adequately considered.

Objective: The aim of our study was to investigate the effects of age, education, and sex on the MoCA total score and to provide demographically adjusted normative values for a German-speaking population.

Methods: Subjects were recruited from a registry of healthy volunteers. Cognitive health was defined using the Mini-Mental State (score ≥ 27/30 points) and the Consortium to Establish a Registry for Alzheimer's Disease-Neuropsychological Assessment Battery (total score ≥ 85.9 points). Participants were assessed with the German version of the MoCA. Normative values were developed based on regression analysis. Covariates were chosen using the Predicted Residual Sums of Squares approach.

Results: The final sample consisted of 283 participants (155 women, 128 men; mean (SD) age = 73.8 (5.2) years; education = 13.6 (2.9) years). Thirty-one percent of participants scored below the original cut-off (< 26/30 points). The MoCA total score was best predicted by a regression model with age, education, and sex as covariates. Older age, lower education, and male sex were associated with a lower MoCA total score ($p < 0.001$).

Conclusion: We developed a formula to provide demographically adjusted standard scores for the MoCA in a German-speaking population. A comparison with other MoCA normative studies revealed considerable differences with respect to selection of volunteers and methods used to establish normative data.

Keywords: Elderly individuals, healthy participants, mild cognitive impairment, Montreal Cognitive Assessment, regression analysis

3.1 INTRODUCTION

Due to the demographical development, age-related diseases will drastically increase over the next decades. Today, 46.7 million people are suffering from dementia worldwide – a number that is estimated to nearly triple by 2050 and reach 131.5 million cases (Prince et al., 2016). To face this healthcare challenge, early and accurate identification of cognitive impairment is crucial. Mild cognitive impairment (MCI) may represent a stage along the clinical continuum of Alzheimer's disease, and currently there are no drugs proven effective for this disease stage (Petersen et al., 2017). However, implementing off-label pharmacological treatment might be beneficial in certain patients; non-pharmacological interventions should be initiated; behavioral or psychiatric symptoms common in MCI may be treated; and there is time to consider important life choices when a patient is still able to do so (Petersen et al., 2017). Additionally, future pharmacological interventions against Alzheimer's disease (AD) mainly target patients in an incipient disease stage (Scheltens et al., 2016), and about 10% of the causes of cognitive impairment are reversible (Clarfield, 2003).

The early detection of cognitive decline requires a tool that is short, easy to administer and interpret, and has high diagnostic accuracy. Currently, a widely-used instrument is the Mini-Mental State Examination (MMSE; Folstein et al., 1975). However, the MMSE sensitivity is poor when identifying individuals with MCI (Ciesielska et al., 2016; Nasreddine et al., 2005; Roalf et al., 2013), and it lacks meaningful assessment of executive functions (Fu et al., 2017). The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) has been developed to address these weaknesses. It has demonstrated better diagnostic accuracy in patients with MCI (Ozer et al., 2016; Trzepacz et al., 2015), has less ceiling effect (Trzepacz et al., 2015), and a higher test-retest-reliability (Ozer et al., 2016). In addition, the MoCA better captures the cognitive domains proposed in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; *American Psychiatric Association,* 2013). Accordingly, previous research has demonstrated good practical utility of the MoCA as a diagnostic tool in various diseases affecting cognition (Borland et al., 2017).

Yet, the implementation of the MoCA has some limitations. First, the proposed cut-off score of 26 out of 30 points (Nasreddine et al., 2005) has been criticized for being too conservative. A recent review found that MoCA specificity was 60% or

lower when applying this cut-off score (Davis et al., 2015), thus, bearing a high risk of false-positive classifications. Second, possible demographic effects on cognitive performance are not well addressed in the original MoCA, which only includes a basic correction for education (+1 point for individuals with ≤ 12 years of education). However, it has been shown that age and – less consistently – sex may influence MoCA scores (Abou-Mrad et al., 2017; Borland et al., 2017; Conti, Bonazzi, Laiacona, Masina, & Coralli, 2015; Freitas, Simoes, Alves, & Santana, 2011; Kenny et al., 2013; Konstantopoulos, Vogazianos, & Doskas, 2016; Kopecek et al., 2017; Larouche et al., 2016; Lu et al., 2011; Malek-Ahmadi et al., 2015; Narazaki et al., 2013; Nasreddine, Phillips, & Chertkow, 2012; Nasreddine et al., 2005; Pereiro et al., 2017; Rossetti, Lacritz, Cullum, & Weiner, 2011; Rossetti et al., 2017; Santangelo et al., 2015). Finally, the MoCA performance may vary across different cultures and languages (Rossetti et al., 2011). Accordingly, normative values for the MoCA have been established in several countries (Abou-Mrad et al., 2017; Borland et al., 2017; Conti et al., 2015; Freitas et al., 2011; Kenny et al., 2013; Konstantopoulos et al., 2016; Kopecek et al., 2017; Larouche et al., 2016; Lu et al., 2011; Malek-Ahmadi et al., 2015; Narazaki et al., 2013; Nasreddine et al., 2012; Nasreddine et al., 2005; Pereiro et al., 2017; Rossetti et al., 2011; Rossetti et al., 2017; Santangelo et al., 2015). The results show great variability; most importantly there are substantial differences regarding the empirically derived MoCA cut-off scores (Abou-Mrad et al., 2017; Borland et al., 2017; Conti et al., 2015; Freitas et al., 2011; Kenny et al., 2013; Konstantopoulos et al., 2016; Kopecek et al., 2017; Larouche et al., 2016; Lu et al., 2011; Malek-Ahmadi et al., 2015; Narazaki et al., 2013; Nasreddine et al., 2012; Nasreddine et al., 2005; Pereiro et al., 2017; Rossetti et al., 2011; Rossetti et al., 2017; Santangelo et al., 2015). Consequently, a general cut-off for all populations might not be suitable, and diagnostic accuracy may be improved when a cut-off score is based on culture-specific and demographically adjusted normative values.

To our knowledge, normative values for the German version of the MoCA have not yet been established. The aim of our study was to evaluate the effects of age, education, and sex on the MoCA and to create demographically adjusted norms for the German version. This report also provides a comparison of normative data from other international samples.

3.2 MATERIALS AND METHODS

3.2.1 Participants

Ethical approval for the study (N° EKNZ 2016-00393) was provided by the *Ethikkommission Nordwest- und Zentralschweiz (EKNZ)* on April 12, 2016. The study was performed in respect of the most recent version of the Declaration of Helsinki and was registered on ClinicalTrials.gov (NCT03246269).

Participants were recruited from an existing Registry of Individuals Interested to Participate in Research established by the Memory Clinic, University Center for Medicine of Aging, Felix Platter Hospital in Basel, Switzerland. The detailed study flow chart is shown in Figure 1. The registry was established in 2013 with approval from the local ethics committee (N° EKBB 280/1). Individuals were informed about the registry and the possibility to sign-up by means of newspaper advertisements, television interviews, and public scientific lectures. Each time a study with normal control subjects was initiated at the Memory Clinic, potential participants with the required demographic characteristics (age, education, sex) were identified from the registry and invited to provide information about their medical history by completing a detailed medical questionnaire (see Supplementary Figure 1 for an English translation of the medical questionnaire). At the beginning of the current study in December 2016, the registry consisted of 2,162 individuals. Seven-hundred and ninety-four had previously provided their medical history and were considered during the recruitment process of this study. Four-hundred and eighty-seven individuals remained eligible for telephone screening after applying inclusion and exclusion criteria (see below). During the telephone screening, a further assessment of exclusion criteria was performed, and 153 subjects were excluded. Thus, 334 individuals were assessed between December 2016 and April 2017, and the data of 283 subjects were included in the final analysis (see study flow chart for details).

During the recruitment process, a stratification of sex (female and male) and age (groups: 65–69, 70–74, 75–79, and > 79 years) was applied to obtain age groups with at least 20 women and 20 men each. The aim was to include only cognitively healthy individuals by applying the following criteria. Inclusion criteria were: (1) age ≥ 65 years, (2) education ≥ 7 years, (3) fluent German-speaking, and (4) provided written informed consent. Subjects who met one of the following criteria were excluded: (1) cognitive impairment (i.e., MMSE < 27/30 and/or Consortium to

Establish a Registry for Alzheimer's Disease-Neuropsychological Assessment Battery [CERAD-NAB] < 85.89; Ehrensperger, Berres, Taylor, & Monsch, 2010), any diagnosis of cognitive impairment), (2) diagnosis and/or symptoms of depression (i.e., Geriatric Depression Scale [GDS]; Yesavage & Sheikh, 1986) > 5/15), (3) severe sensory or motor impairment interfering with cognitive testing, (4) serious somatic disease, (5) any disease or events affecting the central nervous system, (6) cerebrovascular disease, (7) current medication with psychoactive drugs except for benzodiazepines, and (8) participation in a cognitive study within the last 3 months (to avoid practice effects).

Fig. 1. Study flow chart.

[1]Based on neuropsychological test results in previous studies and/or individuals with any diagnosis of cognitive impairment.

[2]Based on information provided in the medical questionnaire.

[3]Signs of depression: reported symptoms of depression and/or current diagnosis of depression and/or current psychotherapy for depression.

[4]Severe sensory or motor impairment: any visual or auditory impairment not correctable with (reading) glasses or hearing aids; motor impairment of the upper extremity (e.g., essential tremor, paresis, dyskinesia).

[5]Serious somatic disease (i.e., current chemo- or radiotherapy; severe cardiac, pulmonary, renal, gastrointestinal, or endocrine disease interfering with everyday functioning).

[6]Disease or event affecting the central nervous system (i.e., meningitis, encephalitis, severe traumatic brain injury with loss of consciousness > 5 minutes, intoxication with neurotoxic substances, prior intracranial neurosurgery, general anesthesia within the last three months, previous or current substance addiction [drugs, alcohol, medication]).

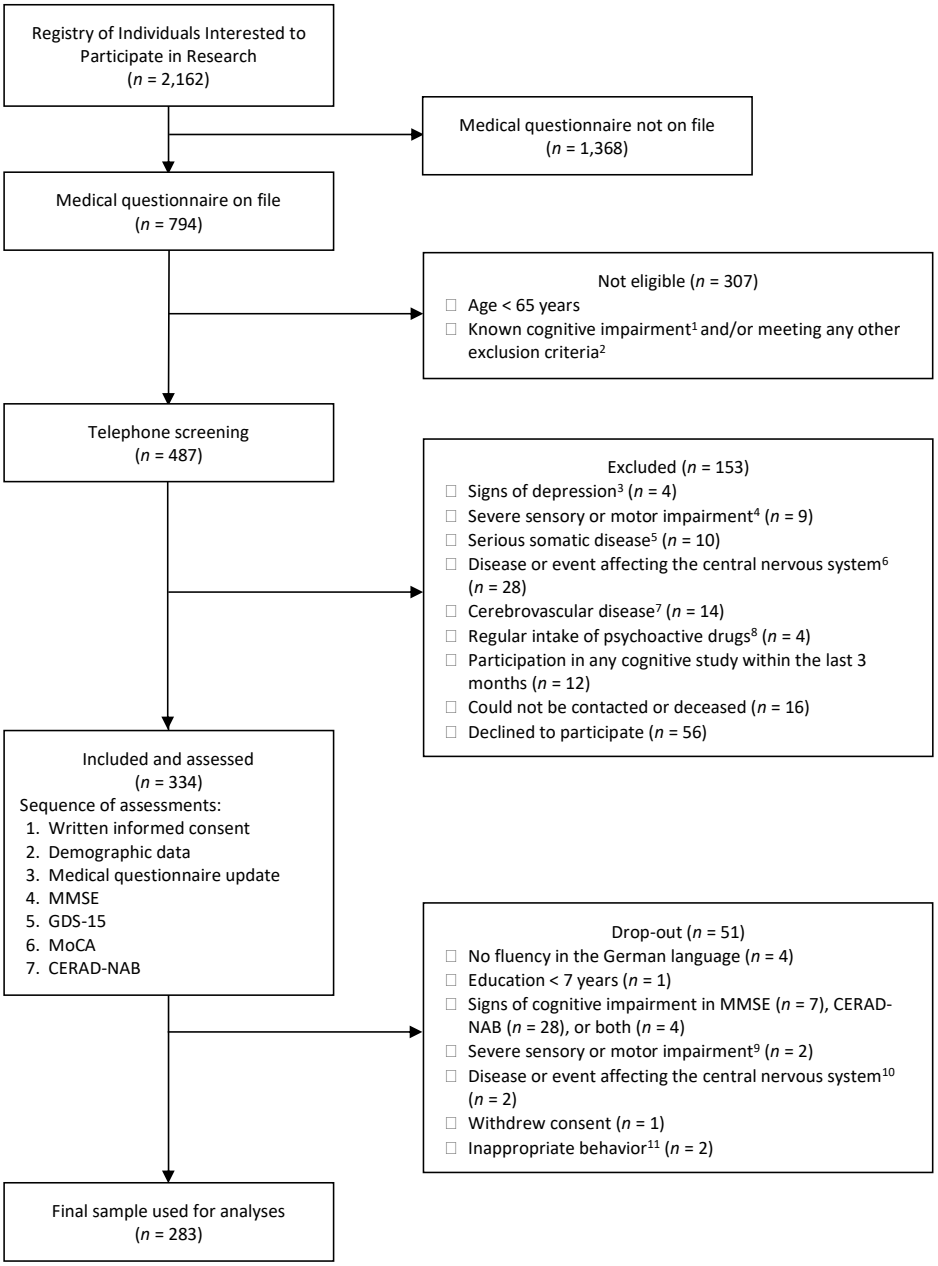
[7]Cerebrovascular disease (i.e., stroke, transient ischemic attack).

[8]Regular intake of psychoactive drugs (i.e., for treatment of schizophrenia, bipolar disorder, obsessive compulsive disorder, personality disorder; substance-induced mental disorder).

[9]Macular degeneration (*n* = 1), hearing impairment interfering with cognitive testing (*n* = 1).

[10]Suspected Parkinson’s disease (*n* = 1), general anesthesia within the last three months (*n* = 1).

Fig. 1 Study Flow chart.



[11]Subject was verbally offensive towards test administrator (*n* = 1); subject deliberately made mistakes during cognitive testing (*n* = 1).

CERAD-NAB = Consortium to Establish a Registry for Alzheimer's Disease-Neuropsychological Assessment Battery; GDS-15 = Geriatric Depression Scale (15 items; no subject scored > 5/15 points); MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment.

3.2.2 Procedures

After obtaining written informed consent, the medical history provided in the medical questionnaire was updated. Then, study eligibility was further assessed with the German versions of the MMSE (Folstein et al., 1975) and the 15-item GDS questionnaire (Yesavage & Sheikh, 1986). After completing these screening procedures, all subjects were assessed with the MoCA. The German version of the CERAD-NAB was administered at the end of the assessment to avoid possible interference effects with the MoCA. The MMSE was neither included in this CERAD-NAB version nor used to calculate the CERAD-NAB total score (Ehrensperger et al., 2010). Subjects meeting any exclusion criteria were omitted from the main statistical analysis only after all assessments took place. One out of four psychology master students who were specifically trained for the study examinations carried out the assessments. All assessments took place on one day during 1-2 hours and were held in a quiet room with subjects seated at a table.

We used the official German translation of the MoCA (Version 7, November 2004; www.mocatest.org). The cognitive domains assessed are: (1) "Visuospatial/Executive", (2) "Naming", (3) "Memory", (4) "Attention", (5) "Language", (6) "Abstraction", (7) "Delayed Recall", and (8) "Orientation". The original version provides an extra point for individuals with lower education (i.e., ≤ 12 years). Since we aimed at diligently correcting for education, we used the uncorrected MoCA total score in our calculations.

3.2.3 Statistical analysis

The effect of age, education, and sex on the MoCA total score was calculated using regression analysis. Twenty different general linear models were tested to adjust for the covariates age, education, and sex. A complete model search between a minimal and a maximal model was performed (Berres, Zehnder, Blasi, & Monsch, 2008). The models included the quantitative covariates, the quantitative covariates' squares, and their interactions with sex (see Supplementary Table 1 for details).

The MoCA total score was transformed using a cubic transformation to achieve normality and homoscedasticity of the residuals. The initial 20 regression models were then recalculated with the transformed score, and the best model was selected. The best model was defined as the model with the minimum Predicted

Residual Sum of Squares (PRESS) statistic. This is a leave-one-out cross-validation with PRESS = $\sum(y_i - \hat{y}_i^{(-i)})^2$ where $\hat{y}_i^{(-i)}$estimates the i^{th} response from a model that was estimated without this observation (Berres et al., 2008). A smaller PRESS statistic indicates a higher predictive power of the corresponding model. The same model was selected before and after transformation, which corroborates the robustness of the method. In a last step, we checked for heterogeneity of variance of the residuals. The formula for the demographically corrected standard scores (z-scores) is based on the final regression model. Normative values were then calculated using the z-score formula.

Sex differences in the MoCA total score were analyzed using the Mann-Whitney U-Test. Spearman's rank correlation for non-parametric data was used to investigate the associations between the MoCA, the CERAD-NAB, and the MMSE total scores. Kendall's Tau for non-parametric data was used to test the associations between the demographic variables and the MoCA subdomains. Raw scores (i.e., not demographically corrected) were used in all analyses.

The required sample size was 171 participants. This allows the estimation of the 5^{th} and the 95^{th} percentile with no more than 2% deviation. Ten additional subjects were included per predictor variable (age, sex, education, and three expected interactions) to account for adjustments in the regression models. Thus, the minimum required sample size was 231 to account for all the predictor variables in the regression model (Jennen-Steinmetz & Wellek, 2005).

All statistical analyses were performed using R, version 3.4.1 (R Foundation, Vienna, Austria) and RStudio Desktop (RStudio, Boston, MA, USA). Data are presented as mean (SD), unless stated otherwise.

3.3 RESULTS

3.3.1 Descriptive analysis

Two hundred and eighty-three cognitively healthy individuals (155 women, 128 men) were included in the final analysis. Participants' mean age was 73.8 (5.2) years, ranging from 65 to 91 years. Education was 13.6 (2.9) years, ranging from 7 to 20 years. The MoCA total score was 26.1 (2.5) points, and the MMSE total score was 29.2 (0.9) points. Detailed demographics are shown in Table 1. Medical

history and current medications of all subjects were assessed based on the medical questionnaire and are displayed in Table 2.

Table 1. Demographic characteristics

Age group	*n*	Age, years	Women, %	Education[1], years	GDS-15 total score	CERAD-NAB total score	MMSE total score	MoCA total score
65–69	68	67.6 (1.4)	61.8	13.2 (2.7)	0.3 (0.8)	97.9 (5.5)	29.4 (0.9)	26.6 (2.6)
70–74	102	72.2 (1.3)	56.9	14.0 (2.9)	0.4 (0.7)	98.6 (5.2)	29.4 (0.7)	26.4 (2.4)
75–79	68	76.5 (1.4)	50.0	13.7 (3.2)	0.3 (0.6)	99.5 (5.9)	29.3 (0.9)	25.8 (2.5)
> 79	45	82.6 (2.4)	46.7	13.3 (2.8)	0.4 (0.7)	99.0 (6.5)	28.9 (1.0)	25.1 (2.4)
Total	283	73.8 (5.2)	54.8	13.6 (2.9)	0.4 (0.7)	98.7 (5.7)	29.2 (0.9)	26.1 (2.5)

Data are presented as mean (SD).

[1]Years of education was defined as the total number of years in school plus any professional education (not counting years needed to repeat). The maximum education was set at 20 years. In case of multiple specialized educations, only the longest one was counted.

CERAD-NAB = Consortium to Establish a Registry for Alzheimer's Disease-Neuropsychological Assessment Battery; GDS-15 = Geriatric Depression Scale (15 items); MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment.

Table 2. Medical history and current medications

Age group	*n*	History of head trauma[1]	Prior general anesthesia[2]	Prior diagnosis of major depr-ession[3]	Prior psychiatric hospitali-zation[4]	Regular alcohol consum-ption[5]	Oral anticoag ulants/ antiplate let drugs	Anti-hyper-tensive drugs	Statins	Oral anti-diabetic drugs
65–69	68	5 (7.4)	59 (86.8)	3 (4.4)	1 (1.5)	45 (66.2)	5 (7.4)	20 (29.4)	14 (20.6)	4 (5.9)
70–74	102	11 (10.8)	85 (83.3)	7 (6.9)	3 (2.9)	61 (59.8)	13 (12.7)	36 (35.3)	14 (13.7)	6 (5.9)
75–79	68	5 (7.4)	56 (82.4)	0 (0)	1 (1.5)	49 (72.1)	17 (25.0)	31 (45.6)	18 (26.5)	1 (1.5)
> 79	45	4 (8.9)	39 (86.7)	1 (2.2)	0 (0)	26 (57.8)	19 (42.2)	25 (55.6)	14 (31.1)	3 (6.7)
Total	283	25 (8.8)	239 (84.5)	11 (3.9)	5 (1.7)	181 (64.0)	54 (19.1)	112 (39.6)	60 (21.2)	14 (4.9)

Data are presented as *n* (%).

[1]Mild head trauma with or without loss of consciousness < 5 minutes.

[2]General anesthesia at least three months prior to study participation.

[3]No current diagnosis of major depression and/or current psychotherapy for major depression.

[4]Due to psychiatric diseases that occurred in the past (e.g., major depression).

[5]Participants answering the question: “Do you drink alcohol regularly?” with: “yes”.

The MoCA total scores ranged from 15 to 30 points when corrected for education (Nasreddine et al., 2005). Their distribution is shown in Figure 2. Eighty-eight of the 283 subjects (31.1%) scored below the cut-off score of < 26/30 points. The mean MoCA total score was higher for women than for men (26.3 (2.4) vs. 25.7 (2.6) points, $p = 0.042$). The rates of subjects with the maximum scores in subdomains were: "Visuospatial/Executive" = 50.2%, "Naming" = 99.3%, "Attention" = 76.0%, "Language" = 52.7%, "Abstraction" = 56.9%, "Delayed Recall" = 29.7%, and "Orientation" = 93.3%. The MoCA total score showed a moderate positive correlation with the CERAD-NAB total score ($r_s = 0.45$, $p < 0.001$) and a weak positive correlation with the MMSE total score ($r_s = 0.20$ $p < 0.001$). A weak positive correlation was also observed between MMSE and CERAD-NAB total scores ($r_s = 0.23$, $p < 0.001$). There were no missing values in any of the analyses.

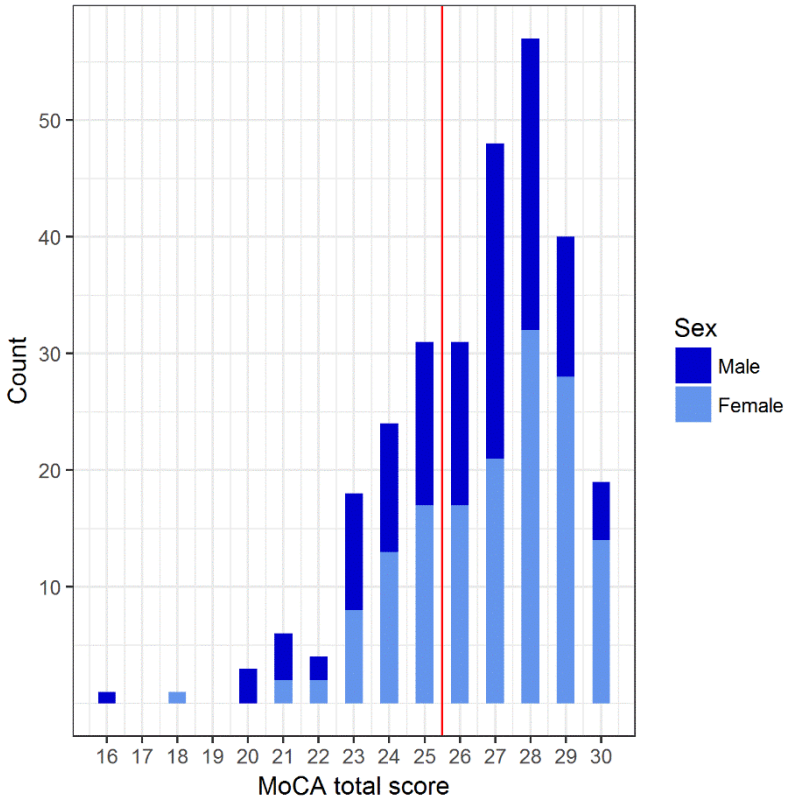


Fig. 2. Distribution of corrected MoCA total scores. The red line indicates the originally proposed MoCA cut-off (26/30 points). In our study, 88 subjects (31.1%) scored below this cut-off.

3.3.2 Demographic influences on the MoCA total score

The MoCA total score was best predicted by a regression model with age, education, and sex (adjusted $R^2 = 0.12$, $F = 14.2$, $p < 0.001$), explaining 12% of the variance. In the regression analysis, increasing age ($p < 0.001$), less education ($p < 0.001$), and male sex ($p = 0.003$) were associated with a lower MoCA total score. The t-values indicate that this effect is strongest for education (t = 4.99), followed by age (t = -3.41), and sex (t = 3.02). The associations between the MoCA total score and demographic characteristics are shown in Figure 3. An analysis of the influence of demographic variables on the MoCA subdomains is presented in Supplementary Table 2.

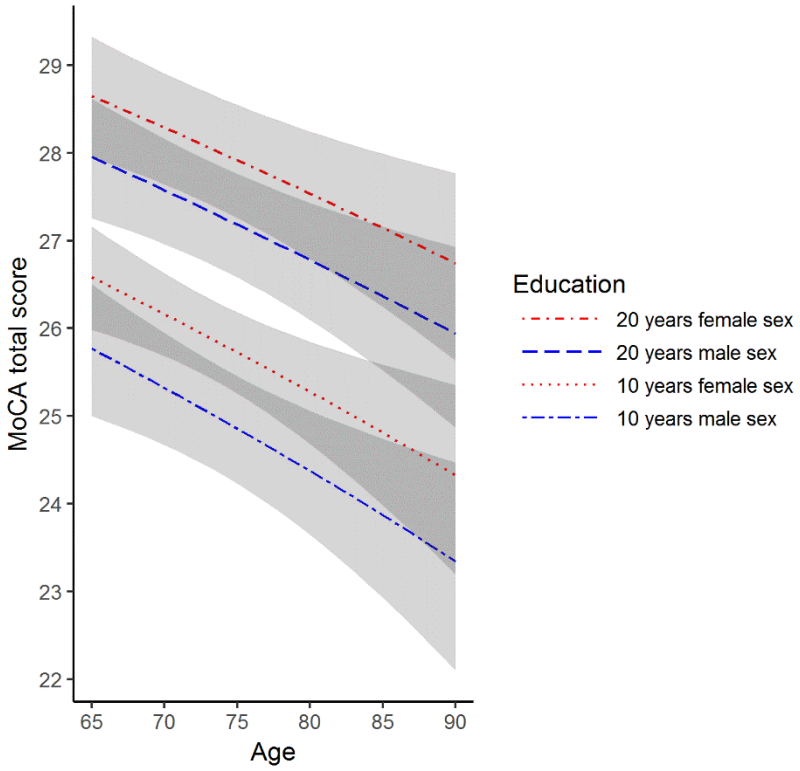


Fig. 3. Association of the MoCA total score with age, education, and sex. Exemplary regression lines are shown for 10 and 20 years of education, respectively. The regression model indicates that the MoCA total score is lower with increasing age and fewer years of education. Overall, female sex was associated with a higher MoCA total score than male sex. The areas in grey represent the 95% confidence intervals.

3.3.3 Z-score calculation

The z-scores are based on the formula: z = (transformed score - expected score) / residual standard deviation. A nearly normal distribution of the residuals was achieved using a cubic transformation of the raw MoCA total score. The formula for the demographically corrected z-score was derived from the final regression model. The z-score can be calculated as follows: z = $MoCA\ total\ score^3$ - (23816.36 + (-175.821 ∗ *age*) + (472.9053 ∗ *education*) + (1672.542 ∗ *sex*)) / 4470.258. Sex is coded as male = 0 and female = 1. Age and education are entered in integer values (years). We followed the example of Weintraub et al. (2018) and will provide a web-based calculation tool (www.mocatest.ch) to automatically determine the z-score by entering the individual demographic data and MoCA total score.

3.3.4 Cut-off scores

Cut-off values were calculated based on the z-score formula (Table 3). The calculation was done separately for women and men for each year of age (65–91) and year of education (7–20). The cut-off was set at a z-score of ≤ -1.28 (10^{th} percentile) to achieve 90% specificity. The applied percentiles may vary depending on the specific setting (e.g., screening in research or case-finding). We, therefore, chose to establish normative tables for the most common percentiles used. All cut-off score tables (i.e., -1.64 SD [5^{th} percentile], -1 SD [16^{th} percentile], -1.5 SD [7^{th} percentile], and -2 SD [2.5^{th} percentile]) are provided in Supplementary Tables 3-6.

Table 3. Highest MoCA total scores located just below the 10th percentile (z-score < -1.28)

Women

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	22	22	23	23	23	24	24	24	24	25	25	25	25	26
66	22	22	23	23	23	24	24	24	24	25	25	25	25	26
67	22	22	23	23	23	23	24	24	24	24	25	25	25	25
68	22	22	22	23	23	23	24	24	24	24	25	25	25	25
69	22	22	22	23	23	23	23	24	24	24	25	25	25	25
70	22	22	22	23	23	23	23	24	24	24	24	25	25	25
71	21	22	22	22	23	23	23	24	24	24	24	25	25	25
72	21	22	22	22	23	23	23	23	24	24	24	24	25	25
73	21	22	22	22	22	23	23	23	24	24	24	24	25	25
74	21	21	22	22	22	23	23	23	24	24	24	24	25	25
75	21	21	22	22	22	23	23	23	23	24	24	24	24	25
76	21	21	22	22	22	22	23	23	23	24	24	24	24	25
77	21	21	21	22	22	22	23	23	23	23	24	24	24	25
78	21	21	21	22	22	22	23	23	23	23	24	24	24	24
79	20	21	21	21	22	22	22	23	23	23	24	24	24	24
80	20	21	21	21	22	22	22	23	23	23	23	24	24	24
81	20	21	21	21	22	22	22	22	23	23	23	24	24	24
82	20	20	21	21	21	22	22	22	23	23	23	24	24	24
83	20	20	21	21	21	22	22	22	23	23	23	23	24	24
84	20	20	20	21	21	22	22	22	22	23	23	23	24	24
85	20	20	20	21	21	21	22	22	22	23	23	23	23	24
86	19	20	20	21	21	21	22	22	22	23	23	23	23	24
87	19	20	20	20	21	21	21	22	22	22	23	23	23	24
88	19	20	20	20	21	21	21	22	22	22	23	23	23	23
89	19	19	20	20	21	21	21	22	22	22	22	23	23	23
90	19	19	20	20	20	21	21	21	22	22	22	23	23	23
91	19	19	20	20	20	21	21	21	22	22	22	23	23	23

Men

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	21	21	22	22	22	23	23	23	23	24	24	24	25	25
66	21	21	22	22	22	22	23	23	23	24	24	24	24	25
67	21	21	21	22	22	22	23	23	23	24	24	24	24	25
68	21	21	21	22	22	22	23	23	23	23	24	24	24	24
69	21	21	21	22	22	22	22	23	23	23	24	24	24	24
70	20	21	21	21	22	22	22	23	23	23	24	24	24	24
71	20	21	21	21	22	22	22	23	23	23	23	24	24	24
72	20	20	21	21	21	22	22	22	23	23	23	24	24	24
73	20	20	21	21	21	22	22	22	23	23	23	23	24	24
74	20	20	21	21	21	22	22	22	23	23	23	23	24	24
75	20	20	20	21	21	21	22	22	22	23	23	23	24	24
76	20	20	20	21	21	21	22	22	22	23	23	23	23	24
77	19	20	20	21	21	21	22	22	22	22	23	23	23	24
78	19	20	20	20	21	21	21	22	22	22	23	23	23	24
79	19	19	20	20	21	21	21	22	22	22	23	23	23	23
80	19	19	20	20	20	21	21	22	22	22	22	23	23	23
81	19	19	20	20	20	21	21	21	22	22	22	23	23	23
82	19	19	19	20	20	21	21	21	22	22	22	23	23	23
83	18	19	19	20	20	20	21	21	21	22	22	22	23	23
84	18	19	19	20	20	20	21	21	21	22	22	22	23	23
85	18	19	19	19	20	20	21	21	21	22	22	22	22	23
86	18	18	19	19	20	20	20	21	21	21	22	22	22	23
87	18	18	19	19	19	20	20	21	21	21	22	22	22	23
88	18	18	19	19	19	20	20	20	21	21	22	22	22	22
89	17	18	18	19	19	20	20	20	21	21	21	22	22	22
90	17	18	18	19	19	19	20	20	21	21	21	22	22	22
91	17	18	18	18	19	19	20	20	20	21	21	21	22	22

The values correspond to the highest raw scores just below the 10th percentile. For instance, a MoCA total score of 22 points is just below the 10th percentile for a 65-year-old woman with 7 years of education.

Note: The bonus point for individuals with ≤ 12 years of education must not be applied when using this cut-off score table.

3.4 DISCUSSION

Our study provides demographically corrected normative values (z-scores) for the German version of the MoCA. The MoCA total score was influenced by age, education, and sex, which is in line with previous normative studies of the MoCA (Borland et al., 2017; Konstantopoulos et al., 2016; Larouche et al., 2016). Other studies found significant effects of age and education, but not for sex (Conti et al., 2015; Freitas et al., 2011; Kenny et al., 2013; Kopecek et al., 2017; Lu et al., 2011; Malek-Ahmadi et al., 2015; Narazaki et al., 2013; Pereiro et al., 2017; Rossetti et al., 2011; Rossetti et al., 2017; Santangelo et al., 2015). While there is a basic adjustment for education in the original version (+ 1 point for education ≤ 12 years), our analyses provide a more precise correction for this important influencing factor. Moreover, we made necessary adjustments for age and sex, which are lacking in the original version.

Considering these demographic influences will likely improve the diagnostic accuracy of the MoCA. For instance, in our sample of cognitively healthy participants, 88 subjects (31.1%) scored below the originally proposed cut-off score of 26 points (Nasreddine et al., 2005), even when the bonus point was given for individuals with ≤ 12 years of education. The demographically corrected cut-off values provided in our study may reduce this false-positive rate. For example, a MoCA total score of 23 in an 85-year-old man (hypothetical patient 1) with 8 years of education is considered to be pathological according to the originally recommended cut-off score, even if one point would be added due to education ≤ 12 years. However, his demographically corrected z-score (based on our study) is -0.11, which is still considered to be within normal limits. In contrast, a MoCA total score of 26 points in a 65-year-old woman (hypothetical patient 2) with 20 years of education is considered to be within normal limits. Yet, her demographically corrected z-score (based on our study) is -1.33, which is below the 10th percentile and, therefore, pathological. These two examples illustrate that using demographically adjusted normative values lead to a decrease of false-positive (hypothetical patient 1) and false-negative results (hypothetical patient 2), respectively.

In our analysis, 12% of the variance in the MoCA total score was explained by demographic characteristics, while other authors reported an explained variance up to 49% (Freitas et al., 2011). This discrepancy is likely due to the much larger age range in some studies. Because both age and education influence cognitive performance, the variance increases when age or education ranges are broad. Consequently, including these variables in a regression model will explain more of the variance. When paralleling our findings to a study with a smaller age range (Borland et al., 2017), results are very comparable ($R^2 = 0.11$).

In our study, the correlation between the MoCA and CERAD-NAB total scores was much higher than the correlation between the MMSE and CERAD-NAB total scores. This suggests that the MoCA assesses cognition in a more comprehensive way compared to the MMSE. Twenty-eight excluded subjects scored below the cut-off on the CERAD-NAB, but still had an MMSE score ≥ 27 points, supporting the notion that the MMSE lacks sensitivity for detection of MCI. In this context, a recent report by Chapman et al. (2016) indicates that the MMSE might be unsuitable to define eligibility for AD clinical trials. There is a clear need for a cognitive screening tool with high diagnostic accuracy for subject enrollment in AD studies. Future

studies may verify whether the MoCA (used with appropriate norms) is more suitable to determine subject selection.

3.4.1 Comparison with international normative samples

In recent years, several research groups conducted normative studies for the MoCA in different languages. An overview of the existing literature is provided in Table 4. The majority of these reports suggest that the originally proposed MoCA cut-off score of 26 points is too conservative. Nine out of 14 normative studies reported a mean MoCA total score < 26 points in their sample (Abou-Mrad et al., 2017; Conti et al., 2015; Freitas et al., 2011; Kopecek et al., 2017; Malek-Ahmadi et al., 2015; Narazaki et al., 2013; Rossetti et al., 2011; Rossetti et al., 2017; Santangelo et al., 2015). In general, studies reported the mean MoCA total score without the one-point correction for education; one study did not mention whether the correction was applied (Freitas et al., 2011). When applying the bonus point for education, nearly one-third of our sample scored below the cut-off of 26 points. Previous normative studies using the original cut-off score reported false-positive rates of 46% (Malek-Ahmadi et al., 2015) up to 76% (Rossetti et al., 2017).

There are several explanations for these high false-positive rates and their substantial variation between studies. First, the MoCA total score might be influenced by intercultural and language differences (e.g., socioeconomic or sociodemographic factors, different word lengths originating from translations [Kopecek et al., 2017; Lu et al., 2011; Rossetti et al., 2017; Strauss, Sherman, & Spreen, 2006]). One study suggests that ethnicity may influence the MoCA total score (Rossetti et al., 2017). However, this may be explained by disparities in socioeconomic factors (e.g., quality of education) rather than ethnicity itself (Strauss et al., 2006). Second, there are important differences in sample sizes, ranging from n = 90 (Nasreddine et al., 2005) to n = 6,283 (Lu et al., 2011). Larger samples may better represent the general population and decrease the risk of sampling errors (Strauss et al., 2006). Yet, even large studies may have small cell sizes, when distinct subgroups (e.g., age categories) are defined to create norms. Third, not all studies were intended as normative studies, and data may have been collected for other purposes (Borland et al., 2017; Larouche et al., 2016; Malek-Ahmadi et al., 2015; Narazaki et al., 2013; Pereiro et al., 2017; Rossetti et al., 2011; Rossetti et al., 2017). These "samples of convenience" may lack appropriate

inclusion/exclusion criteria and standard procedures in MoCA administration, leading to increased variability within samples, especially if data are gathered from multiple centers (Strauss et al., 2006). Fourth, there are substantial dissimilarities in the demographic characteristics of study participants. Mean age differs by almost 40 years between the youngest (Konstantopoulos et al., 2016) and the oldest sample (Malek-Ahmadi et al., 2015). Large variances can also be seen in mean education, ranging from 8.2 (4.7) (Freitas et al., 2011) to 14.4 (3.8) (Larouche et al., 2016) years. Considering the effects of these demographic characteristics on the MoCA performance, differences in mean age or education possibly lead to variances in the mean MoCA total score among studies. Finally, normative studies diverge regarding inclusion/exclusion criteria (Borland et al., 2017; Narazaki et al., 2013; Nasreddine et al., 2012). Cognitive health of participants is of utmost importance in normative studies, particularly if subtle cognitive changes should be detected. In some normative studies, cognition was assessed using methods that might not be sensitive enough to detect subtle cognitive impairment (Kenny et al., 2013; Rossetti et al., 2011; Rossetti et al., 2017). Other investigators did not screen for cognitive impairment at all (Narazaki et al., 2013).

3.4.2 Cognitive health in normative samples

There are two different methodological approaches to normative studies. One is to rely on a population-based sample to create norms; for the other, a sample of indisputably healthy volunteers is chosen. Both methods bear the risk of inducing bias: while the former is prone to false-negative errors, the latter is prone to false-positive ones (Strauss et al., 2006). In our study, we chose the latter approach and applied stringent criteria to assure cognitive health of the participants. One might argue that such rigorous exclusion criteria may lead to a sample of "supernormal" individuals. However, the population-based approach does not seem appropriate when normative data are collected for an elderly population. Since the incidence and prevalence of MCI increases with age (Petersen et al., 2017), the probability of erroneously including individuals suffering from a cognitive disorder increases as well. Including cognitively impaired individuals in a normative group lowers the reference range for cognitive health, and the distinction between the two groups (MCI vs. healthy individuals) will be less clear. Consequently, it is very likely that the sensitivity for the detection of MCI decreases when relying on a population-based approach. Thus, we consider the criteria of indisputable cognitive health as a mandatory prerequisite for normative data.

Table 4. Overview of international normative data for the MoCA

Source publication	MoCA language	Study type	*n*	Age, years	Age range, years	Female, %	Education, years	Exclusion of medical comorbidities[1]	Exclusion of cognitive impairment	Method	Output	Demographic effects	MMSE score	MoCA total score (uncorrected[2])	MoCA total score (corrected[2])
Abou-Mradi et al. (2017)[39]	Arabic	NS	164	70.1 (6.9)	60–87	58.5	n.r.	Yes	Yes (IQCODE)	Frequency analysis, regression models	Base-rates for the 5th, 10th, 15th percentiles, z-scores, regression equation	Age and education not significant)	28.6 (1.3)	24.2 (2.9)	n.r.
Borland et al. (2017)[38]	Swedish	NS	758	73.1 (5.1)	65–85	62.5	n.r.	Yes	Yes (MMSE ≥ 24 and AQT ≥ 290 s, and some subjects: neuropsychological assessment, brain imaging, CSF analysis)	Regression models, intercept, estimates, RMSE	Percentiles, z-scores, regression equation	Age, Education	27.9 (1.4)	26.0 (2.3)	n.r.
Contin et al. (2015)[41]	Italian	NS	225	70.1 (5.7)	60–80	50.7	9.9 (4.6)	Yes	Yes (adjusted MMSE ≥ 23.8 and adjusted PMT ≤ 5.25)	Regression models	Correction grid for total scores, regression formula	Age, Education	n.r.	23.3 (3.2)	74*
Freitas et al. (2011)[40]	Portuguese	NS	650	55.8 (15.1)	25–91	62.8	8.2 (4.7)	Yes	Yes (abnormal performance in neuropsychological assessment, CDR)	Regression models	Mean (SD) cut-off scores for 2 SD, 1.5 SDs, and 1 SDs below the mean	Age, Education	28.9 (1.3)[3]	24.7 (3.7)*	n.r.
Kenny et al. (2013)[37]	English	NS	5802	63.1 (n.r.)	n.r.	54.7	n.r.	n.r.	Yes (MMSE ≥ 10, diagnosis of dementia, AD or PD)	GAMLSS	Percentiles, mean (SD)	Age, Education	n.r.	n.r.	n.r.

Konstantopoulos et al. (2016)	Greek	NS	710	46.9 (16.6)	20–85	53.8	13.8 (3.8)	Yes	Yes (some questions of CDR, > 1.5 SDs below the normative values in > 1 neuropsychological tests)	Regression models	Percentiles, mean (SD) for the total and subscores	A, E, S	n. p.	27.2 (1.9)*	n. r.
Kopecek et al. (2017)	Czech	NS	540	75.6 (9.1)	n. r.	54.1	12.7 (3.5)	Yes	Yes (2 SDs below the sample mean in > 2 neuropsychological tests)	Regression models	Percentiles	A, E	27.4 (2.0)[3]	24.7 (2.9)[5]	50
Larouche et al. (2016)	Quebec-French	NS	1019	67.8 (8.8)	n. r.	67.3	14.4 (3.8)	Yes	Yes (neuropsychological assessment in most subjects, self-report in all subjects)	Regression models, PRESS, cross-validation	z-scores, regression equation	A, E, S	n. p.	26.4 (2.7)	n. r.
Lu et al. (2011)	Chinese	VS	6283	72.0 (0.8)	65–100	52.1	6.7 (1.1)	Yes	Yes (CDR > 0)	n. r.	Mean (SD)	A, E	26.3 (0.6)	23.8 (0.9)*	48
Malek-Ahmadi et al. (2015)	English	NS	205	84.7 (7.9)	70–99	68.3	n. r.	Yes	Yes (MMSE < 26)	ANOVA	Mean (SD)	A, E	28.9 (1.2)	25.0 (3.1)	46
Narazaki et al. (2013)	Japanese	NS	1977	73.6 (6.2)	65–96	58.7	11.0 (2.5)	Yes	No	Regression models	Mean (SD), median, range	A, E	n. p.	21.8 (3.9)	75
Nasreddine et al. (2005)	French/ English	VS	90 (NC)	72.8 (7.0)	n. r.	60.0	13.3 (3.4)	Only in 51 subjects	Yes (neuropsychological assessment)	n. r.	Cut-off score	E	n. r.	26.9 (n. r.)*[4]	13[5]

Pereiro et al. (2017)	Spanish	NS	563	66.43 (11.2)	n. r.	57.2	9.83 (4.8)	Yes	Yes (MMSE > 1.5 SD below the mean of the corresponding age and education group)	Regression coefficient, SD of the residuals	Percentiles, regression equation	A, E	27.93 (2.1)	n. r.	n. r.
Rossetti et al. (2011)	English	NS	2653	50.33 (11.2)	18–85	60.0	13.43 (2.5)	Only stroke	Yes (3 yes/no questions regarding subjective cognitive impairment)	Pearson correlation, analysis of variance	Mean (SD)	A, E	n. p.	23.43 (4.0)	62
Rossetti et al. (2017)	English	NS	1118	n. r.	n. r.	n. r.	n. r.	Only stroke	Yes (3 yes/no questions regarding subjective cognitive impairment)	Pearson correlation, independent samples t-test	Mean (SD), median	A, E	n. p.	22.33 (3.9)	76
Santangelo et al. (2015)	Italian	NS	415	56.8 (18.8)	21–95	60.7	11.1 (4.8)	Yes	Yes (MMSE below the demographically corrected cut-off score)	Regression models	Correction grid for total scores, cut-off score	A, E	28.0 (2.4)	22.0 (4.2)	n. r.
Current study	German	NS	283	73.8 (5.2)	65–91	54.8	13.6 (2.9)	Yes	Yes (MMSE < 27 + CERAD total score < 85.89)	Regression models, PRESS	Percentiles, z-scores, regression equation	A, E, S	29.2 (0.9)	26.1 (2.5)	31

Studies are ordered alphabetically. Data are presented as mean (SD), unless stated otherwise.

[1]Medical comorbidities (e.g., neurologic, cerebrovascular and/or psychiatric diseases known to affect cognition).

[2]Studies marked with an asterisk (*) did not indicate if the presented MoCA total score is corrected for education or not.

[3]Pooled mean (SD).

[4]As reported by Nasreddine et al. (2012).

[5]As reported by Abou-Mrad et al. (2017), based on the reported specificity of 0.87 by Nasreddine et al. (2005).

A = age; AD = Alzheimer's disease; AQT = A Quick Test of Cognitive Speed; ANOVA = analysis of variance; DQ = Dementia Questionnaire; CDR = Clinical Dementia Rating; c. s. = corrected score; CSF = cerebrospinal fluid; E = education; GAMLSS = Generalized Additive Models for Location, Shape, and Scale; S = sex; MoCA = Montreal Cognitive Assessment; MMSE = Mini-Mental State Examination; n. r. = not reported; n. p. = not performed; NC = normal controls; NS = normative study; PD = Parkinson's disease; PMT = Prose Memory Test; VS = validation study (with control group); PRESS = Predicted Residual Sum of Squares; RMSE = Root Mean Square Error; SD = standard deviation.

3.4.3 Strengths and limitations

A regression-based approach yields some important advantages over the traditional norming method (i.e., reference ranges for cells of age and/or education groups). First, in traditional norming the sample is divided into subgroups. This leads to relatively small sample sizes per group, even if the overall sample size is quite large (Berres et al., 2008). In contrast, regression-based norming considers the whole sample, and the continuous variables (i.e., age and education) are analyzed in their full range. Second, relying on age and/or education groups to create norms may misrepresent individuals who are situated close to the boundary of a subgroup (Larouche et al., 2016). Moreover, due to the more or less arbitrarily chosen subgroup boundaries, traditional norming may not properly reflect the natural development of cognitive performance (Strauss et al., 2006). The regression-based approach, however, considers the overall trend in the data. Third, the regression-based approach allows to simultaneously study multiple covariates and their potential interactions.

We acknowledge some limitations of our study. First, there may be a selection bias as our participants were recruited from an existing registry of individuals interested in taking part in research projects. These individuals may potentially show a greater motivation to perform well in cognitive testing than the average population. Individuals who participated in this study completed the Swiss educational system. Although the educational system in Switzerland is not 100% equal to the educational systems in other German-speaking countries, we believe that the acquired normative data are suitable for German-speaking populations in general. Our norms are intended for the elderly population and cannot be applied to individuals younger than 65 years. Second, cognitive test performance is commonly adjusted for demographic influences. Yet, some authors question if demographic adjustments are appropriate in dementia diagnostics, because age and education are known risk factors for cognitive impairment (Narazaki et al., 2013; Strauss et al., 2006). O'Connell and Tuokko (2010) found that the overall diagnostic accuracy is comparable for raw versus adjusted scores. While having lower sensitivity, the adjusted scores were shown to have better specificity. As our results show, MoCA performance declines with older age and/or lower education (Table 2). Therefore, when using a simple cut-off, the rate of false-positives may be higher with increasing age and/or lower education. Thus, adjusted scores may be more appropriate if the MoCA is used for diagnostic purposes in elderly individuals.

Our aim was to enhance the sensitivity of the MoCA by excluding any individuals with signs of cognitive impairment. In addition, specificity likely increases when applying a demographic adjustment of the obtained total score. However, the current normative data are not suitable to determine the exact diagnostic accuracy of the German MoCA. This version of the MoCA must first be validated in cognitively impaired patients, which is a follow-up project.

3.5 CONCLUSIONS

This study provides normative values for the German version of the MoCA. Our findings support the frequent statement that the originally proposed cut-off score may be too conservative. The MoCA performance was influenced by age, education and, – less consistently – by sex in all available studies, including ours. Thus, using demographically adjusted norms will improve the diagnostic accuracy of the MoCA. In addition, we observed a high level of heterogeneity in the methodology of existing normative studies. Therefore, we strongly suggest an international harmonization of guidelines for normative studies to enhance comparability in the future.

ACKNOWLEDGEMENTS

We gratefully acknowledge the help of all participating volunteers, as well as the staff of the Memory Clinic, University Center for Medicine of Aging Basel, Felix Platter Hospital, Basel, Switzerland. In particular, we thank Panagiota Mistridis, PhD, and Sabine Krumm, PhD, for providing advice during study conduction, and Ursi Kunze, MSc, for her support in database management. We thank Andrina Baitella, BSc, Sarah Roesiger, BSc, Jael Fasnacht, BSc, and Antonino Cusano, BSc, for their help in conducting the study interviews and Allison Dwileski, BSc, for proof-reading the manuscript.

Sources of support

Financial support for this research was provided from internal sources from the Memory Clinic, University Center for Medicine of Aging, Felix Platter Hospital, Basel, Switzerland and from the Department of Anesthesia, Surgical Intensive Care, Prehospital Emergency Medicine and Pain Therapy, University Hospital Basel, Basel, Switzerland. There was no external financial and material support.

Conflicts of interest

The authors have no conflict of interest to report.

Supplementary material

Supplementary Table 1. Twenty regression models

Basic models	+ square of age	+ square of education	+ square of age + square of education
A + E	A + E + A^2	A + E + E^2	A + E + A^2 + E^2
A + E + S	A + E + S + A^2	A + E + S + E^2	A + E + S + A^2 + E^2
A + E + S + AS	A + E + S + AS + A^2	A + E + S + AS + E^2	A + E + S + AS + A^2 + E^2
A + E + S + ES	A + E + S + ES + A^2	A + E + S + ES + E^2	A + E + S + ES + A^2 + E^2
A + E + S + AS + ES	A + E + S + AS + ES + A^2	A + E + S + AS + ES + E^2	A + E + S + AS + ES + A^2 + E^2

A = age; AS = (age - mean age) * sex; A^2 = (age - mean age)^2; E = education; ES = (education - mean education) * sex; E^2 = (education - mean education)^2; S = sex.

Supplementary Table 2. Correlations (Kendall's Tau) between demographical variables and MoCA subdomains

	Age	Education	Female sex
Visuospatial/Executive	-0.11*	0.22***	-0.16*
Naming	0.00	0.07	-0.08
Attention	-0.08	0.10	-0.03
Language	-0.02	0.15**	0.15*
Abstraction	-0.03	0.17***	0.04
Delayed Recall	-0.12*	0.03	0.17**
Orientation	0.03	-0.03	0.04
MoCA total score	-0.15***	0.19***	0.11*

* $p < 0.05$; ** $p < 0.005$; *** $p < 0.001$

Supplementary Table 3. Highest MoCA total scores located just below the 5th percentile

Women

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	21	21	22	22	22	23	23	23	23	24	24	24	25	25
66	21	21	22	22	22	23	23	23	23	24	24	24	24	25
67	21	21	21	22	22	22	23	23	23	24	24	24	24	25
68	21	21	21	22	22	22	23	23	23	23	24	24	24	25
69	21	21	21	22	22	22	23	23	23	23	24	24	24	24
70	20	21	21	21	22	22	22	23	23	23	24	24	24	24
71	20	21	21	21	22	22	22	23	23	23	23	24	24	24
72	20	21	21	21	22	22	22	22	23	23	23	24	24	24
73	20	20	21	21	21	22	22	22	23	23	23	24	24	24
74	20	20	21	21	21	22	22	22	23	23	23	23	24	24
75	20	20	20	21	21	21	22	22	22	23	23	23	24	24
76	20	20	20	21	21	21	22	22	22	23	23	23	23	24
77	19	20	20	21	21	21	22	22	22	23	23	23	23	24
78	19	20	20	20	21	21	21	22	22	22	23	23	23	24
79	19	20	20	20	21	21	21	22	22	22	23	23	23	23
80	19	19	20	20	21	21	21	22	22	22	22	23	23	23
81	19	19	20	20	20	21	21	21	22	22	22	23	23	23
82	19	19	19	20	20	21	21	21	22	22	22	23	23	23
83	19	19	19	20	20	20	21	21	22	22	22	22	23	23
84	18	19	19	20	20	20	21	21	21	22	22	22	23	23
85	18	19	19	19	20	20	21	21	21	22	22	22	23	23
86	18	18	19	19	20	20	20	21	21	21	22	22	22	23
87	18	18	19	19	20	20	20	21	21	21	22	22	22	23
88	18	18	19	19	19	20	20	21	21	21	22	22	22	22
89	17	18	18	19	19	20	20	20	21	21	21	22	22	22
90	17	18	18	19	19	20	20	20	21	21	21	22	22	22
91	17	18	18	19	19	19	20	20	20	21	21	22	22	22

Men

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	20	20	21	21	21	22	22	22	22	23	23	23	24	24
66	20	20	20	21	21	21	22	22	22	23	23	23	24	24
67	20	20	20	21	21	21	22	22	22	23	23	23	23	24
68	19	20	20	20	21	21	22	22	22	22	23	23	23	24
69	19	20	20	20	21	21	21	22	22	22	23	23	23	23
70	19	19	20	20	21	21	21	22	22	22	23	23	23	23
71	19	19	20	20	20	21	21	21	22	22	22	23	23	23
72	19	19	20	20	20	21	21	21	22	22	22	23	23	23
73	19	19	19	20	20	21	21	21	22	22	22	22	23	23
74	18	19	19	20	20	20	21	21	21	22	22	22	23	23
75	18	19	19	20	20	20	21	21	21	22	22	22	23	23
76	18	19	19	19	20	20	21	21	21	22	22	22	22	23
77	18	18	19	19	20	20	20	21	21	21	22	22	22	23
78	18	18	19	19	19	20	20	21	21	21	22	22	22	23
79	18	18	18	19	19	20	20	20	21	21	21	22	22	22
80	17	18	18	19	19	20	20	20	21	21	21	22	22	22
81	17	18	18	19	19	19	20	20	21	21	21	22	22	22
82	17	18	18	18	19	19	20	20	20	21	21	21	22	22
83	17	17	18	18	19	19	20	20	20	21	21	21	22	22
84	17	17	18	18	19	19	19	20	20	21	21	21	22	22
85	16	17	17	18	18	19	19	20	20	20	21	21	21	22
86	16	17	17	18	18	19	19	19	20	20	21	21	21	22
87	16	17	17	18	18	19	19	19	20	20	20	21	21	21
88	16	16	17	17	18	18	19	19	20	20	20	21	21	21
89	16	16	17	17	18	18	19	19	19	20	20	21	21	21
90	15	16	17	17	18	18	18	19	19	20	20	20	21	21
91	15	16	16	17	17	18	18	19	19	20	20	20	21	21

The values presented correspond to the highest raw scores just below the 5th percentile. For instance, a MoCA total score of 21 points is just below the 5th percentile for a 65-year-old woman with 7 years of education. Note: The bonus point for individuals with ≤ 12 years of education must not be applied when using this normative table.

Supplementary Table 4. Highest MoCA total scores located just below -1 SD (16th percentile)

Women

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	23	23	24	24	24	24	25	25	25	25	26	26	26	26
66	23	23	23	24	24	24	24	25	25	25	25	26	26	26
67	23	23	23	24	24	24	24	25	25	25	25	26	26	26
68	23	23	23	23	24	24	24	25	25	25	25	25	26	26
69	23	23	23	23	24	24	24	24	25	25	25	25	26	26
70	22	23	23	23	24	24	24	24	25	25	25	25	26	26
71	22	23	23	23	23	24	24	24	25	25	25	25	25	26
72	22	22	23	23	23	24	24	24	24	25	25	25	25	26
73	22	22	23	23	23	24	24	24	24	25	25	25	25	26
74	22	22	23	23	23	23	24	24	24	24	25	25	25	25
75	22	22	22	23	23	23	24	24	24	24	25	25	25	25
76	22	22	22	23	23	23	23	24	24	24	25	25	25	25
77	22	22	22	23	23	23	23	24	24	24	24	25	25	25
78	21	22	22	22	23	23	23	24	24	24	24	25	25	25
79	21	22	22	22	23	23	23	23	24	24	24	25	25	25
80	21	22	22	22	22	23	23	23	24	24	24	24	25	25
81	21	21	22	22	22	23	23	23	24	24	24	24	25	25
82	21	21	22	22	22	23	23	23	23	24	24	24	24	25
83	21	21	22	22	22	22	23	23	23	24	24	24	24	25
84	21	21	21	22	22	22	23	23	23	23	24	24	24	25
85	21	21	21	22	22	22	23	23	23	23	24	24	24	24
86	20	21	21	21	22	22	22	23	23	23	24	24	24	24
87	20	21	21	21	22	22	22	23	23	23	23	24	24	24
88	20	21	21	21	22	22	22	23	23	23	23	24	24	24
89	20	20	21	21	21	22	22	22	23	23	23	24	24	24
90	20	20	21	21	21	22	22	22	23	23	23	24	24	24
91	20	20	21	21	21	22	22	22	22	23	23	23	24	24

Men

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	22	22	23	23	23	23	24	24	24	24	25	25	25	25
66	22	22	22	23	23	23	24	24	24	24	25	25	25	25
67	22	22	22	23	23	23	23	24	24	24	24	25	25	25
68	22	22	22	22	23	23	23	24	24	24	24	25	25	25
69	21	22	22	22	23	23	23	24	24	24	24	25	25	25
70	21	22	22	22	23	23	23	23	24	24	24	24	25	25
71	21	21	22	22	22	23	23	23	24	24	24	24	25	25
72	21	21	22	22	22	23	23	23	23	24	24	24	25	25
73	21	21	22	22	22	23	23	23	23	24	24	24	24	25
74	21	21	21	22	22	22	23	23	23	24	24	24	24	25
75	21	21	21	22	22	22	23	23	23	23	24	24	24	24
76	21	21	21	22	22	22	23	23	23	24	24	24	24	24
77	20	21	21	21	22	22	22	23	23	23	24	24	24	24
78	20	21	21	21	22	22	22	23	23	23	23	24	24	24
79	20	20	21	21	22	22	22	22	23	23	23	24	24	24
80	20	20	21	21	21	22	22	22	23	23	23	23	24	24
81	20	20	21	21	21	22	22	22	23	23	23	23	24	24
82	20	20	20	21	21	21	22	22	22	23	23	23	24	24
83	20	20	20	21	21	21	22	22	22	23	23	23	23	24
84	19	20	20	21	21	21	22	22	22	22	23	23	23	24
85	19	20	20	20	21	21	21	22	22	22	23	23	23	24
86	19	20	20	20	21	21	21	22	22	22	23	23	23	23
87	19	19	20	20	20	21	21	22	22	22	22	23	23	23
88	19	19	20	20	20	21	21	21	22	22	22	23	23	23
89	19	19	19	20	20	21	21	21	22	22	22	23	23	23
90	18	19	19	20	20	20	21	21	21	22	22	22	23	23
91	18	19	19	20	20	20	21	21	21	22	22	22	23	23

The values presented correspond to the highest raw scores just below -1 SD. For instance, a MoCA total score of 23 points is just below -1 SD for a 65-year-old woman with 7 years of education. Note: The bonus point for individuals with ≤ 12 years of education must not be applied when using this normative table.

Supplementary Table 5. Highest MoCA total scores located just below -1.5 SD (7th percentile)

Women

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	22	22	22	22	23	23	23	24	24	24	24	25	25	25
66	21	22	22	22	23	23	23	23	24	24	24	25	25	25
67	21	22	22	22	23	23	23	23	24	24	24	24	25	25
68	21	21	22	22	22	23	23	23	24	24	24	24	25	25
69	21	21	22	22	22	23	23	23	23	24	24	24	25	25
70	21	21	22	22	22	22	23	23	23	24	24	24	24	25
71	21	21	21	22	22	22	23	23	23	24	24	24	24	25
72	21	21	21	22	22	22	23	23	23	23	24	24	24	24
73	20	21	21	22	22	22	22	23	23	23	24	24	24	24
74	20	21	21	21	22	22	22	23	23	23	23	24	24	24
75	20	21	21	21	22	22	22	23	23	23	23	24	24	24
76	20	20	21	21	21	22	22	22	23	23	23	24	24	24
77	20	20	21	21	21	22	22	22	23	23	23	23	24	24
78	20	20	21	21	21	22	22	22	22	23	23	23	24	24
79	20	20	20	21	21	21	22	22	22	23	23	23	24	24
80	20	20	20	21	21	21	22	22	22	23	23	23	23	24
81	19	20	20	21	21	21	22	22	22	22	23	23	23	24
82	19	20	20	20	21	21	21	22	22	22	23	23	23	24
83	19	19	20	20	21	21	21	22	22	22	23	23	23	23
84	19	19	20	20	20	21	21	21	22	22	22	23	23	23
85	19	19	20	20	20	21	21	21	22	22	22	23	23	23
86	19	19	19	20	20	21	21	21	22	22	22	23	23	23
87	18	19	19	20	20	20	21	21	21	22	22	22	23	23
88	18	19	19	20	20	20	21	21	21	22	22	22	23	23
89	18	19	19	19	20	20	21	21	21	22	22	22	22	23
90	18	18	19	19	20	20	20	21	21	21	22	22	22	23
91	18	18	19	19	19	20	20	21	21	21	22	22	22	23

Men

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	20	21	21	21	22	22	22	23	23	23	23	24	24	24
66	20	21	21	21	22	22	22	22	23	23	23	24	24	24
67	20	20	21	21	21	22	22	22	23	23	23	24	24	24
68	20	20	21	21	21	22	22	22	23	23	23	23	24	24
69	20	20	20	21	21	22	22	22	22	23	23	23	24	24
70	20	20	20	21	21	21	22	22	22	23	23	23	23	24
71	19	20	20	21	21	21	22	22	22	23	23	23	23	24
72	19	20	20	20	21	21	21	22	22	22	23	23	23	24
73	19	20	20	20	21	21	21	22	22	22	23	23	23	23
74	19	19	20	20	21	21	21	22	22	22	22	23	23	23
75	19	19	20	20	20	21	21	21	22	22	22	23	23	23
76	19	19	20	20	20	21	21	21	22	22	22	23	23	23
77	19	19	19	20	20	20	21	21	22	22	22	22	23	23
78	18	19	19	20	20	20	21	21	21	22	22	22	23	23
79	18	19	19	19	20	20	21	21	21	22	22	22	23	23
80	18	18	19	19	20	20	20	21	21	21	22	22	22	23
81	18	18	19	19	20	20	20	21	21	21	22	22	22	23
82	18	18	19	19	19	20	20	21	21	21	22	22	22	22
83	17	18	18	19	19	20	20	20	21	21	21	22	22	22
84	17	18	18	19	19	20	20	20	21	21	21	22	22	22
85	17	18	18	19	19	19	20	20	20	21	21	22	22	22
86	17	17	18	18	19	19	20	20	20	21	21	21	22	22
87	17	17	18	18	19	19	19	20	20	21	21	21	22	22
88	17	17	18	18	18	19	19	20	20	20	21	21	21	22
89	16	17	17	18	18	19	19	20	20	20	21	21	21	22
90	16	17	17	18	18	19	19	19	20	20	21	21	21	22
91	16	16	17	18	18	18	19	19	20	20	20	21	21	21

The values presented correspond to the highest raw scores just below -1.5 SD. For instance, a MoCA total score of 22 points is just below -1.5 SD for a 65-year-old woman with 7 years of education. Note: The bonus point for individuals with ≤ 12 years of education must not be applied when using this normative table.

Supplementary Table 6. Highest MoCA total scores located just below -2 SD (2.5th percentile)

Women

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	20	20	21	21	21	22	22	22	23	23	23	23	24	24
66	20	20	20	21	21	21	22	22	22	23	23	23	24	24
67	20	20	20	21	21	21	22	22	22	23	23	23	23	24
68	19	20	20	21	21	21	22	22	22	22	23	23	23	24
69	19	20	20	20	21	21	21	22	22	22	23	23	23	24
70	19	20	20	20	21	21	21	22	22	22	23	23	23	23
71	19	19	20	20	21	21	21	22	22	22	22	23	23	23
72	19	19	20	20	20	21	21	21	22	22	22	23	23	23
73	19	19	19	20	20	21	21	21	22	22	22	23	23	23
74	18	19	19	20	20	20	21	21	21	22	22	22	23	23
75	18	19	19	20	20	20	21	21	21	22	22	22	23	23
76	18	19	19	19	20	20	21	21	21	22	22	22	23	23
77	18	18	19	19	20	20	20	21	21	21	22	22	22	23
78	18	18	19	19	20	20	20	21	21	21	22	22	22	23
79	18	18	19	19	19	20	20	21	21	21	22	22	22	22
80	17	18	18	19	19	20	20	20	21	21	21	22	22	22
81	17	18	18	19	19	19	20	20	21	21	21	22	22	22
82	17	18	18	19	19	19	20	20	20	21	21	21	22	22
83	17	17	18	18	19	19	20	20	20	21	21	21	22	22
84	17	17	18	18	19	19	19	20	20	21	21	21	22	22
85	17	17	18	18	18	19	19	20	20	20	21	21	21	22
86	16	17	17	18	18	19	19	20	20	20	21	21	21	22
87	16	17	17	18	18	19	19	19	20	20	21	21	21	22
88	16	16	17	17	18	18	19	19	20	20	20	21	21	21
89	16	16	17	17	18	18	19	19	19	20	20	21	21	21
90	15	16	17	17	18	18	19	19	19	20	20	20	21	21
91	15	16	16	17	17	18	18	19	19	20	20	20	21	21

Men

Age (years) \ Education (years)	7	8	9	10	11	12	13	14	15	16	17	18	19	20
65	18	19	19	20	20	20	21	21	21	22	22	22	23	23
66	18	19	19	20	20	20	21	21	21	22	22	22	23	23
67	18	19	19	19	20	20	20	21	21	22	22	22	22	23
68	18	18	19	19	20	20	20	21	21	21	22	22	22	23
69	18	18	19	19	19	20	20	21	21	21	22	22	22	23
70	18	18	18	19	19	20	20	20	21	21	21	22	22	22
71	17	18	18	19	19	20	20	20	21	21	21	22	22	22
72	17	18	18	19	19	19	20	20	21	21	21	22	22	22
73	17	17	18	18	19	19	20	20	20	21	21	21	22	22
74	17	17	18	18	19	19	20	20	20	21	21	21	22	22
75	17	17	18	18	19	19	19	20	20	20	21	21	22	22
76	16	17	17	18	18	19	19	20	20	20	21	21	21	22
77	16	17	17	18	18	19	19	19	20	20	21	21	21	22
78	16	17	17	18	18	18	19	19	20	20	20	21	21	21
79	16	16	17	17	18	18	19	19	20	20	20	21	21	21
80	16	16	17	17	18	18	19	19	19	20	20	21	21	21
81	15	16	16	17	18	18	18	19	19	20	20	20	21	21
82	15	16	16	17	17	18	18	19	19	20	20	20	21	21
83	15	15	16	17	17	18	18	19	19	19	20	20	21	21
84	15	15	16	16	17	17	18	18	19	19	20	20	20	21
85	14	15	16	16	17	17	18	18	19	19	19	20	20	21
86	14	15	15	16	17	17	18	18	18	19	19	20	20	20
87	14	14	15	16	16	17	17	18	18	19	19	20	20	20
88	13	14	15	16	16	17	17	18	18	19	19	19	20	20
89	13	14	15	15	16	16	17	18	18	18	19	19	20	20
90	13	14	14	15	16	16	17	17	18	18	19	19	20	20
91	12	13	14	15	15	16	17	17	18	18	19	19	19	20

The values presented correspond to the highest raw scores just below -2 SD. For instance, a MoCA total score of 20 points is just below -2 SD for a 65-year-old woman with 7 years of education. Note: The bonus point for individuals with ≤ 12 years of education must not be applied when using this normative table.

Supplementary Fig. 1. English translation of the medical questionnaire.

Geriatric Psychiatry
Memory Clinic

felixplatterhospital
University Center for Medicine of Aging

Questionnaire for Individuals Interested to Participate in Research

ID-No.:

1. Do you currently feel healthy? ☐Yes ☐No

   .....

2. Are you currently suffering from any disease? If yes, please specify: ☐Yes ☐No
   (e.g., diabetes, macular degeneration, etc.)

   .....

   .....

   .....

3. Did you suffer from any serious illness in the past? ☐Yes ☐No
   If yes, please specify: **Age**

   ....

   ....

   ....

   ....

4. Have you ever undergone surgery? If yes, please specify: **Age** ☐Yes ☐No

   ....

   ....

   ....

   ....

5. Have you had any serious accidents? If yes, please specify: **Age** ☐Yes ☐No

   ....

   ....

6. Have you ever had a head injury (possibly with unconsciousness)? ☐Yes ☐No

7. Are you currently taking any medication? If yes, please specify: ☐Yes ☐No

   Name of the medication *Reason*

   .....

   .....

   .....

   .....

8.	Can you hear well? Are you able to easily follow a conversation (with hearing aids if needed)?	☐Yes ☐No
9.	Are you able to read the newspaper (if necessary with glasses or contact lenses, but without magnifying glass)?	☐Yes ☐No
10.	Did you recently notice any trouble remembering new things?	☐Yes ☐No
11.	When did you have your last general anesthesia? Year/Month: ...../.....	☐Never
12.	Have you ever been diagnosed with major depression?	☐Yes ☐No
13.	Did you ever undergo inpatient psychiatric treatment?	☐Yes ☐No
14.	Do you drink alcohol regularly? If yes, how much? (e.g., 1-2 glasses of wine on weekends)	☐Yes ☐No
15.	Have you ever consumed mind-altering substances (e.g., cannabis, marijuana, LSD, cocaine, crack)?	☐Yes ☐No
16.	Is there any metal inside your body (e.g., artificial joint)?	☐Yes ☐No
17.	Are you having difficulties in confined spaces (claustrophobia)?	☐Yes ☐No
18.	Do you have blood relatives with brain disorders (e.g., memory impairment)?	☐Yes ☐No

Additional comments:

.....

.....

.....

I understand that by completing this questionnaire, I will not automatically take part in a study. My details are stored in an electronic database. These data are kept absolutely confidential and are only accessible to the responsible person within the Memory Clinic. Data will not be passed on to third parties. I have the right, without giving reasons, to withdraw my registration at any time. In this case, all collected data will be deleted.

Place, date _____

Signature _____

4. Study II:

Two separate cut-offs on the MoCA for patients with a neurocognitive disorder

Alessandra E. Thomann[a,b], Manfred Berres[c], Nicolai Goettel[b,d], Luzius A. Steiner[b,d], and Andreas U. Monsch[a]

Submitted

[a]Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel, Switzerland
[b]Anesthesiology, University Hospital Basel, Switzerland
[c]Department of Mathematics and Technology, University of Applied Sciences Koblenz, Koblenz, Germany
[d]Department of Clinical Research, University of Basel, Switzerland

ABSTRACT

Introduction: The Montreal Cognitive Assessment (MoCA) has good sensitivity for mild cognitive impairment, but specificity is low when the original cut-off is used. We aim to revise the cut-off on the German MoCA for its use in clinical routine.

Methods: Data were analyzed from 496 Memory Clinic outpatients (447 individuals with a neurocognitive disorder; 49 with cognitive normal findings) and from 283 normal controls. Cut-offs were identified based on (1) the Youden's index and (2) the 10th percentile of the control group.

Results: Compared to the original, a cut-off of 23/24 points had higher specificity (92% vs 63%), but lower sensitivity (65% vs 86%). Introducing two separate cut-offs increased diagnostic accuracies with 92% specificity (23/24 points) and 91% sensitivity (26/27 points). Scores between these two cut-offs require further examinations.

Discussion: Using two separate cut-offs for the MoCA combined with scores in a gray area enhances the accuracy of cognitive screening.

Keywords: Sensitivity and Specificity; Neuropsychology; Mental Status and Dementia Tests; Montreal Cognitive Assessment; Mini Mental State Examination; Neurocognitive Disorders; ROC Curve; Cognitive Dysfunction; Area Under Curve

4.1 BACKGROUND

A steep increase in the prevalence of dementia is expected (Prince et al., 2016) and early detection of cognitive decline is crucial (Yaffe, 2018). The implementation of therapeutic strategies depends on a successful case-finding process at the general practitioners' office and reliable screening tools are required (Ehrensperger et al., 2014). In addition, accurate cognitive assessment allows for an adequate selection of participants in clinical research, since erroneous inclusion or exclusion of individuals may bias study findings (Edmonds et al., 2018; Edmonds et al., 2016).

The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) has gained popularity for cognitive screening. It correlates well with extensive neuropsychological test batteries (Lam et al., 2013; Thomann, Goettel, Monsch, et al., 2018) and covers most of the cognitive domains outlined in the Diagnostic and Statistical Manual, 5th Edition (DSM-5; *American Psychiatric Association,* 2013). However, while the initially proposed cut-off (25/26 points; Nasreddine et al., 2005) has shown good sensitivity for mild cognitive impairment (MCI) (Ozer et al., 2016; Trzepacz et al., 2015), this cut-off may lead to an unacceptably high number of false-positive classifications (Carson et al., 2018; Davis et al., 2015; Thomann, Goettel, Monsch, et al., 2018). Consequently, new cut-offs have been proposed for various patient populations and languages (see Carson et al., 2018 for an overview).

The properties of any screening test are not fixed characteristics, but depend on the clinical context (Florkowski, 2008), limiting the transferability of findings to other settings. Moreover, optimal cut-offs are likely to be specific to the individual study (Martin, Schroeder, & Baade, 2017; Weissberger et al., 2017) and should be validated in independent samples. In clinical practice, patient populations are typically heterogeneous, and medical comorbidities are frequent. In most validation studies, a rather homogenous patient sample was recruited (e.g., only patients with probable Alzheimer's disease [AD] according to McKhann-Criteria; McKhann et al., 2011), exclusion of patients with medical comorbidities). Excluding patients who are difficult to diagnose induces several forms of bias and may lead to an overestimation of diagnostic accuracy (Lijmer et al., 1999; Noel-Storr et al., 2014). Heterogeneous samples reflect the clinical reality more accurately as health care professionals face the challenge to identify truly impaired patients from a pool of

individuals with a suspected neurocognitive disorder (NCD), irrespective of its underlying cause.

In the present study, we estimate the diagnostic accuracy of the original MoCA cut-off in consecutive Memory Clinic outpatients (MC sample) to differentiate normal findings (NF; i.e., neurocognitive results were within normal limits) from patients with mild and major NCD (labelled Mild+Major NCD in the following). Since the MoCA was developed to identify individuals with MCI, subgroup analyses are performed for patients diagnosed with mild NCD (labelled Mild NCD in the following). Given the high rate of false-positive classifications that is associated with the original MoCA cut-off, we aimed at finding a new cut-off with higher specificity. In this context, we introduce a novel approach to determine a cut-off solely based on a sample of cognitively healthy normal controls (NC), which is then validated in the MC sample. In sub-analyses, we investigate the differences in diagnostic accuracy in relation to demographic adjustments by comparing the original MoCA score with recently established demographically corrected MoCA z-scores (Thomann, Goettel, Monsch, et al., 2018). During our analyses, we noticed that information is lost when a continuous variable like the MoCA is dichotomized (M. D. Brown & Reeves, 2003; Vermeersch et al., 2009), and a traditional binary cut-off is used. We therefore propose a new approach to evaluate cognitive performance on the MoCA using two separate cut-offs in combination with a gray area between these scores.

4.2 METHODS

4.2.1 Participants

We retrospectively assessed data from 1,307 consecutive outpatients of the Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel, Switzerland, undergoing neuropsychological assessment between March 6, 2017 and October 12, 2018. Data from patients meeting the following inclusion criteria were considered for the analysis: (1) age ≥65 years, (2) education ≥7 years, (3) fluency in the German language, and (4) availability of a neuropsychological assessment including the Mini Mental Status Examination (MMSE; Folstein et al., 1975) and the MoCA. Exclusion criteria were: (1) severe sensory or motor impairment interfering with cognitive testing, (2) repeated testing with the MoCA

due to follow-up examinations, and (3) documented refusal of the use of personal health-related data for research purposes. An overview of the clinical diagnoses is provided in Supplementary Table 1. The demographic inclusion criteria were selected to match the NC group from a previous normative study on the MoCA (see Thomann, Goettel, Monsch, et al., 2018 for details).

4.2.2 Procedures

Patients were assessed in the following order: (1) detailed patient and medical history, (2) neuropsychological screening including the MMSE and the clock drawing test, (3) MoCA, (4) assessment of symptoms of depression, and (5) extensive neuropsychological examination. Neuropsychological assessments were performed by board-certified neuropsychologists and by psychologists with a Master's degree. Neuropsychological test results were interpreted based on demographically corrected (i.e., age, sex, and education) z-scores. The patients were medically examined by a neurologist or a geriatrician. Imaging (i.e., structural magnetic resonance imaging, computed tomography and/or positron emission tomography with ^{18}F-fluorodeoxyglucose) was performed and in some patients, cerebrospinal fluid was collected to assess for protein deposition. Diagnostic consensus was reached in weekly interdisciplinary meetings by neuropsychologists, neurologists, neuroradiologists, positron emission tomography specialists, geriatricians, psychiatrists, and a neuropathologist. MoCA results were not considered in the diagnostic process. Procedures for the NC group are described elsewhere (Thomann, Goettel, Monsch, et al., 2018).

The study protocol (N° EKNZ 2018-00737) was approved by the regional research ethics board (*Ethikkommission Nordwest- und Zentralschweiz* [EKNZ]) on May 22, 2018. The study was conducted in respect of the most recent version of the Declaration of Helsinki and was registered on ClinicalTrials.gov (NCT03581643). The need for informed consent was waived by the EKNZ.

4.2.3 Statistical analyses

Demographical characteristics and test scores were compared pairwise using the non-parametric Wilcoxon rank sum test for between-group comparisons. Differences in sex were analyzed using the chi-squared test.

Diagnostic accuracies of the original MoCA cut-off (25/26 points) were calculated in the MC sample. Using the Optimal Cutpoints Package in R (López-Ratón, Rodríguez-Álvarez, Cadarso-Suárez, & Gude-Sampedro, 2014), the Youden's index (Sensitivity+Specificity-1; Youden, 1950), was applied to define the optimal cut-offs in the MC sample for the MoCA score, the MoCA z-score, and the MMSE. Additionally, MoCA cut-offs were derived based on the NC group. Specificity was held at approximately 90% in the NC group by choosing the scores that split the sample at the 10^{th} percentile. Hence, normality is defined as a reference range based on the distribution of scores in cognitively healthy individuals, and scores below the 10^{th} percentile were considered pathological. The resulting cut-offs were then validated in the MC sample to differentiate Mild+Major NCD vs. NF and Mild NCD vs. NF.

The discriminative power of the MoCA score, the MoCA z-score, and the MMSE score was estimated in terms of area under the curve (AUC) in the MC sample. Receiver Operating Characteristic (ROC) curves were calculated using the pROC package in R (Robin et al., 2011). The AUCs of the MoCA score vs. MoCA z-score, the MoCA score vs. MMSE score, and the MoCA z-score vs. MMSE score were compared with a bootstrap two-sided significance test for correlated ROC curves. The correct classification rates of the newly derived MoCA cut-offs were compared to the original MoCA cut-off and the optimal cut-offs on the MMSE using the McNemar's test. Results were corrected for multiple comparisons according to Bonferroni-Holm.

We created a plot to visualize the relationship between MoCA scores and rates of sensitivity and specificity. For this purpose, cumulative frequencies were calculated separately for Mild NCD and for NC for each MoCA score. Thus, the proportion of individuals who performed equally or below a given score was determined for each score and expressed in percent of the whole sample. The cumulative frequency for a given score in Mild NCD corresponds to the sensitivity. Specificity is represented by the complementary sum (1 - cumulative frequency) in NC.

All statistical analyses were performed using R, version 3.5.0 (R Foundation, Vienna, Austria) and RStudio Desktop (RStudio, Boston, MA, USA). Data are presented as mean (SD), and the education-corrected MoCA score (+1 point for <12 years of education) was used, unless stated otherwise. There were no missing data in any of the analyses.

4.3 RESULTS

4.3.1 Descriptive analysis

Four hundred and forty-seven patients (Mild+Major NCD), 49 normal findings (NF), and 283 normal controls (NC) were included in the final analysis. Demographic characteristics are displayed in Table 1. There were no differences between NC and NF. Compared to NF, the patients (i.e., Mild+Major NCD, Mild NCD) were older (*P* value < .001), had fewer years of formal education (*P* value < .001), and lower test scores (MMSE: *P* value < .001; MoCA: *P* value < .001, MoCA z-score: *P* value < .001). There were no sex differences between the groups.

Table 1. Demographic characteristics.

Group	*n*	Prevalence in MC sample %	Age (y)	Age range (y)	Education (y)	Education range (y)	Female %	MMSE score	MoCA score	MoCA score range	MoCA z-score	MoCA z-score range
NC	283	-	73.8 (5.2)	65-91	13.6 (2.9)	7-20	54.8	29.2 (0.9)	26.5 (2.4)	16-30	0.0 (1.0)	-3.0-2.4
NF	49	9.9	73.1 (5.6)	65-88	13.8 (2.7)	8-20	40.8	29.0 (1.0)	26.5 (2.2)	22-30	0.1 (1.0)	-1.7-1.9
Mild+Major NCD	447	90.1	78.3 (5.9) *	65-91	12.2 (3.0) *	7-20	55.7	25.1 (3.5) *	19.1 (4.5) *	2-30	-2.1 (1.0) *	-4.3-1.5
Mild NCD	159	32.1	76.0 (6.0) *	65-91	12.4 (3.1) *	7-20	53.5	27.2 (2.2) *	22.0 (3.6) *	12-30	-1.5 (1.0) *	-3.7-1.5

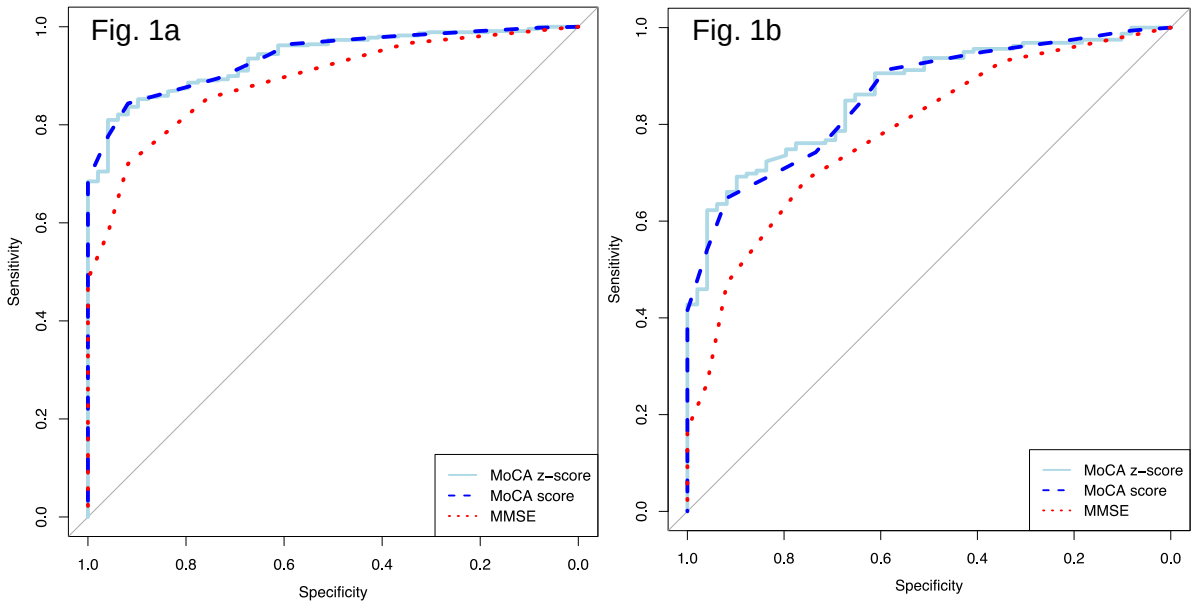
Data are presented as mean (SD). Mild MCD is a subgroup of Mild+Major NCD.

MMSE = Mini Mental-State Examination; MoCA = Montreal Cognitive Assessment; z-score = demographically corrected standard score based on the formula by Thomann, Goettel, Monsch et al. (2018). There were no differences between NC and NF. NF is compared to Mild+Major NCD and Mild NCD: * *P* value < .001.

4.3.2 Diagnostic accuracies

ROC-curves for the MC sample are displayed in Fig. 1. The AUC of the MoCA scores appear larger than that of the MMSE. However, with application of the Bonferroni-Holm procedure, the AUC neither differed significantly between MoCA and MMSE scores (MoCA [AUC=0.94] vs. MMSE [AUC =0.84]: *P* value = .051; MoCA z-score [AUC = 0.94] vs. MMSE: *P* value = .074) nor between the uncorrected MoCA and the MoCA z-score (*P* value = 1.0).

Figure 1. ROC-Curves.



ROC curves for the MoCA (z-score: solid line, corrected score: dashed line) and the MMSE (dotted line) for the classification of Mild+Major NCD (Fig. 1a) and Mild NCD (Fig. 1b).

Cut-offs and the corresponding diagnostic properties for the MoCA and the MMSE are provided in Table 2. Diagnostic accuracies for the MoCA z-score are illustrated in Supplementary Table 2. A MoCA score of 23/24 points was the optimal cut-off according to the 10th percentile-method as well as according to the Youden's index in all patient groups. This cut-off had better correct classification rates than the original MoCA cut-off (25/26 points; *P* value < .001) and the MMSE score (*P* value < .001) in both patient samples. Specificity for the cut-off of 23/24 points was high with 92%, and it had good sensitivity for Mild+Major NCD (84%). However, sensitivity was low for Mild NCD (65%). The original MoCA cut-off (25/26 points) had high sensitivity for Mild+Major NCD (94%) and for Mild NCD (86%), but poor specificity (63%). For Mild NCD, an intermediate cut-off (24/25 points) had neither good sensitivity (74%) nor good specificity (74%). We, therefore, aimed at obviating this trade-off between sensitivity and specificity by defining two separate cut-offs. This is illustrated by the example of Mild NCD vs. NC in section 3.3.

Table 2. Empirically derived cut-offs and diagnostic accuracy estimates for the MoCA and the MMSE.

Measure	**Mild+Major NCD vs. NF**	**Mild NCD vs. NF**
MoCA		
AUC (95% CI, DeLong)	**0.94 (0.91-0.96)**	**0.86 (0.81-0.91)**
Original cut-off	**25/26**	**25/26**
Correct classification rate*	79%	75%
Sensitivity (95% CI)	94% (94-95%)	86% (84-87%)
Specificity (95% CI)	63% (60-67%)	63% (60-67%)
Balanced cut-off	**24/25**	**24/25**
Correct classification rate*	82%	74%
Sensitivity (95% CI)	90% (89-90%)	74% (72-76%)
Specificity (95% CI)	74% (70-77%)	74% (70-76%)
Cut-off; Youden's index†	**23/24**	**23/24**
Cut-off; 10th percentile in NCs	**23/24**	**23/24**
Correct classification rate*	88%	79%
Sensitivity (95% CI)	84% (83-85%)	65% (63-67%)
Specificity (95% CI)	92% (90-94%)	92% (90-94%)
MMSE		
AUC (95% CI, DeLong)	**0.89 (0.85-0.93)**	**0.78 (0.72-0.85)**
Cutoff; Youden's index†	**27/28**	**28/29**
Correct classification rate*	82%	73%
Sensitivity (95% CI)	72% (71-73%)	69% (67-70%)
Specificity (95% CI)	92% (90-94%)	76% (72-79%)

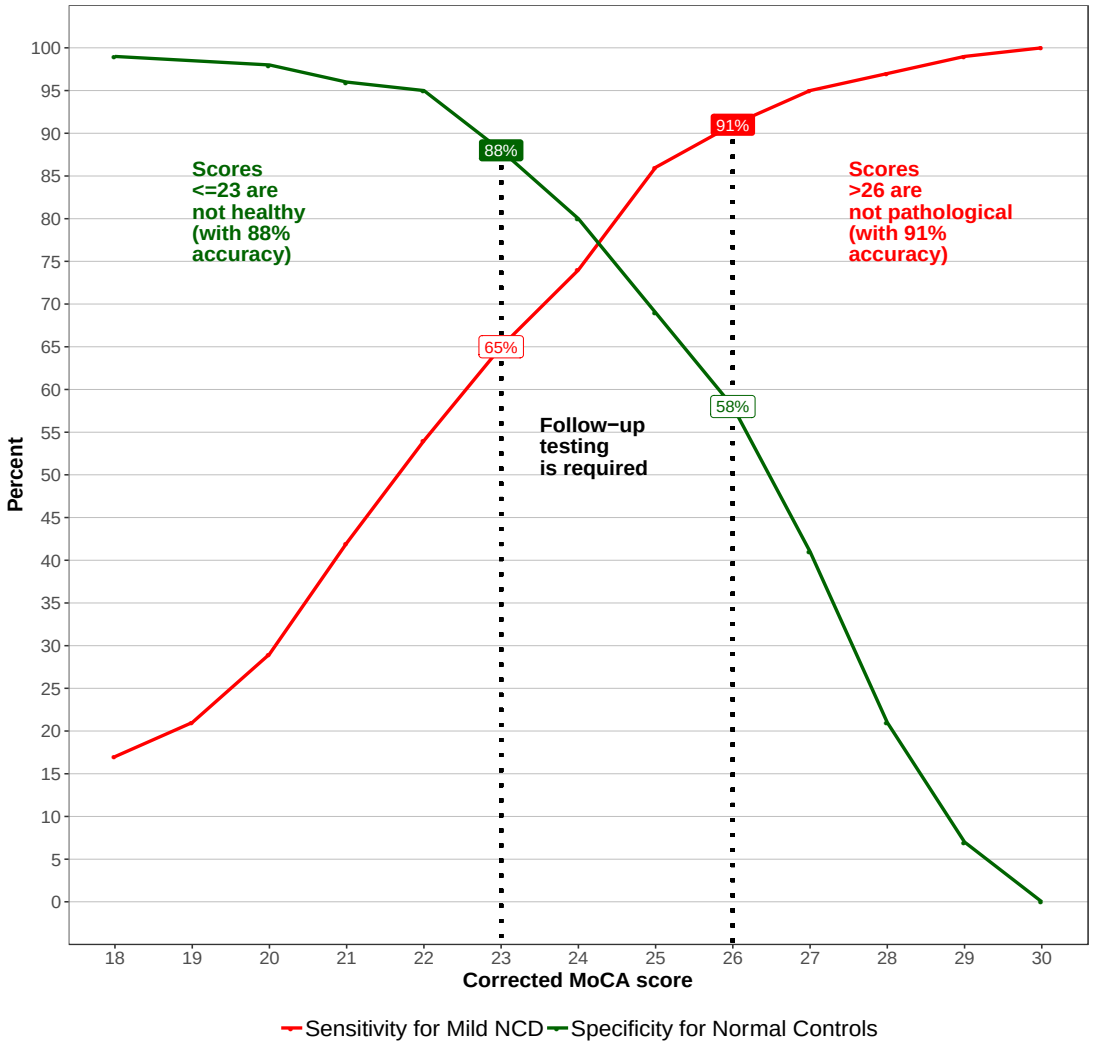
*Correct classification rate = (Sensitivity + Specificity)/2

†Youden's index = Sensitivity + Specificity - 1

4.3.3 Two separate cut-offs and a gray area

In Fig. 2, sensitivity based on Mild NCD is plotted against specificity based on NC. Specificity increases with lower scores, while sensitivity increases with higher scores. At 23/24 points, specificity is 88%, indicating that only 12% of the NC scored ≤23 points. At 26/27 points, sensitivity is 91%, so only 9% of patients with Mild NCD achieved scores >26 points. Consequently, cognitive health and cognitive impairment may be defined using two separate cut-offs.

Figure 2. Two separate cut-offs and a gray area.



The percentage of patients with Mild NCD who were correctly classified as patients (sensitivity, red line) and the percentage of normal controls that were correctly classified as normal controls (specificity, green line) are illustrated. Two cut-offs are illustrated by the dashed lines, one cut-off for not-healthy results (23/24; with 88% specificity) and one cut-off for not-pathological results (26/27; with 91% sensitivity). Scores between these two cut-offs constitute a gray area, where information from further examinations is required.

Analogous to the concept of z-scores, a distribution of scores is assumed, and extreme values are considered improbable for a specific population. For cognitive health, values below a given cut-off (i.e., 23 points) are rare, suggesting that an individual scoring ≤23 points is probably not healthy. This statement was accurate in 88% of the NC group (= specificity). Values above a given cut-off (i.e., >26 points) are uncommon in Mild NCD. Therefore, an individual who attains >26 points on the MoCA probably does not suffer from an NCD. This statement was accurate in 91%

of Mild NCD patients (= sensitivity). Scores between these two cut-offs (24, 25, and 26 points) constitute a gray area. This gray area may be greater or smaller, depending on the desired accuracies (i.e., for sensitivity of 95% and specificity of 95%, the gray area would encompass MoCA scores from 23 to 26 points). The corresponding positive predictive values (PPV) and negative predictive values (NPV) are plotted in Supplementary Fig. 1.

4.4 DISCUSSION

The German MoCA showed good AUC, sensitivity, and specificity for the classification of patients with mild and major NCD versus cognitively healthy normal findings when applied in a heterogeneous group of individuals referred to a university-affiliated Memory Clinic. In the present study, a MoCA score of 23/24 points was established as the optimal cut-off across different patient groups based on two methods. This finding is in line with a recent meta-analysis including seven validation studies on the MoCA (Carson et al., 2018). The new MoCA cut-off had an improved correct classification rate compared to both, the original MoCA cut-off and the MMSE. Further, differences in diagnostic accuracy depending on the severity of cognitive impairment (Mild vs. Major NCD) were revealed. While a cut-off of 23/24 points had high sensitivity for all patients (Mild+Major NCD [84%]), sensitivity was low for Mild NCD (65%). When applying a higher cut-off (e.g., the originally proposed 25/26 points), sensitivity for Mild NCD increased to 91%; however, specificity to detect NF was low (59%). If both measures are balanced, neither of them is sufficiently high. Indeed, most screening tools for MCI lack either sensitivity or specificity (Summers & Bondi, 2017). Since there are currently no effective treatments for most of the underlying causes of MCI, there is no reason to favor one over the other. Based on the findings of this study, we propose a new method to evaluate cognitive performance, taking the MoCA as an example. Instead of applying a single cut-off, two separate cut-offs may be used. One cut-off for results that are unlikely within the normal range, and one cut-off for scores that are rarely seen in patients. MoCA scores >26 points may be considered as not pathological with very high accuracy, while scores ≤23 points are very likely not healthy. Between these scores, we have defined a gray area. When some individual scores within this gray area, the clinician should gather more information to guide

decision-making or perform a follow-up testing in approximately six to twelve months.

4.4.1 Choice of normative samples and patient characteristics

It has been argued that a restrictive cognitively healthy normative group may not be entirely comparable to the population, who is typically screened with the MoCA. This may artificially boost specificity of a test and lead to an overestimation in diagnostic accuracy (Martin et al., 2017; Noel-Storr et al., 2014). We addressed this issue by analyzing two groups of cognitively healthy individuals: one that was purposely recruited for a previous normative study (NC), and one that was formed by consecutively referred patients with a cognitive normal finding (NF). In our study, there were no differences between the NC and the NF group, neither in demographic characteristics nor in cognitive performance. Furthermore, the optimal MoCA cut-offs were identical in these two groups. This suggests that the healthy controls in our study are representative for individuals with cognitive normal findings in the clinical routine. While this is reassuring, longitudinal data from individuals, who remained healthy for several years, should be analyzed in future studies.

4.4.2 Influence of demographic adjustments on diagnostic accuracy

The utility of demographical adjustments has been questioned in previous reports (Strauss et al., 2006), since age and education are *per se* risk factors of cognitive decline. Indeed, patients with mild or major NCD in our study were older and had less years of formal education when compared to the NF group. On the other hand, some authors (including our group) have suggested that correcting for demographical effects may increase diagnostic accuracy when evaluating cognitive performance (Carson et al., 2018; Thomann, Goettel, Monsch, et al., 2018). Conversely, we found no difference between demographically corrected and uncorrected MoCA scores in the overall diagnostic accuracy measured by the AUC. However, a difference emerged in the balance of sensitivity and specificity. When considering the effects of age, education, and sex (z-scores), the MoCA gained specificity, while the uncorrected MoCA score showed increased sensitivity. The education-corrected MoCA score was located in between, with higher sensitivity

but lower specificity compared to the MoCA z-score, and lower sensitivity but higher specificity compared to the uncorrected MoCA score. This result is in line with previous findings from a simulation (O'Connell & Tuokko, 2010). Whether to rely on a demographically adjusted score or on an uncorrected raw score may depend on the setting. For instance, when the MoCA is applied to identify cognitively healthy participants in clinical research, high sensitivity might be more important to avoid the inclusion of patients with false-negative test results. In contrast, high specificity should be favored over sensitivity to avoid including healthy individuals with false-positive results if cognitively impaired patients are included in a clinical trial. Indeed, the erroneous inclusion of cognitively healthy individuals as patients may mask possible treatment effects in clinical trials (Edmonds et al., 2018). When a general practitioner should decide whether to refer a patient to a specialized Memory Clinic based on cognitive screening, false-positive results should be minimized to reduce discomfort for the individual and healthcare costs. On the other hand, false-negative results may deprive a patient of the early implementation of therapeutic strategies. In this situation, we suggest relying on our new system with two separate cut-offs and a gray zone.

4.4.3 Limitations

Sensitivity, specificity, and the AUC give an indication of the quality of the test under observation by classifying the test performance with respect to a reference standard (i.e., an individual will be classified as a patient on the MoCA as well as according to a complete Memory Clinic diagnostic workup). However, these measures do not inform about the probability whether a tested individual has a specific disease (Trevethan, 2017; Weissberger et al., 2017). Predictive values, which are influenced by prevalence rates, reflect this information. In the current study, the MoCA had very high PPV across all patient groups and most MoCA scores (see Supplementary Fig. 1). However, the PPV will be lower in a setting with a low prevalence of disease (e.g., when screening for cognitive impairment at the general practitioner's office). Likewise, in most MoCA studies reporting PPV and NPV, the prevalence of MCI was greater than in the general population (Ozer et al., 2016). Ideally, the diagnostic accuracy of a test should be evaluated in the same setting where it is clinically applied (Habibzadeh, Habibzadeh, & Yadollahie, 2016). We did not have access to any data from first step screening processes (i.e., from a general practitioner's office). Thus, our findings inform about how well the MoCA

classifies individuals as healthy or cognitively impaired compared to a more extensive, multi-dimensional, diagnostic process, as performed in our Memory Clinic (described in 2.2. Procedures). Additionally, we can provide the probability for a Memory Clinic patient to be affected by a mild or major NCD, when the MoCA performance is below the cut-off (PPV), as well as the probability that the patient is cognitively healthy, when the performance lies above the cut-off (NPV; Trevethan, 2017). Therefore, our findings do not inform conclusively about the probability of having mild or major NCD in any other setting than the Memory Clinic. We refer to the excellent recent publication by Trevethan (2017) for a better understanding on the informative value of sensitivity, specificity, PPV, and NPV.

4.5 CONCLUSIONS

In the present study, the diagnostic properties of the German MoCA were evaluated in an outpatient sample referred to a university-affiliated Memory Clinic. The originally proposed MoCA cut-off (25/26 points) had good sensitivity for mild and major NCD, but specificity was poor. As an alternative, a cut-off of 23/24 points on the MoCA improved specificity. However, the sensitivity to detect mild NCD was low using this cut-off. Thus, both cut-offs lead to a trade-off in either sensitivity (23/24 points) or specificity (25/26 points). In this context, we propose a new method to guide clinical decision making by relying on two separate cut-offs combined with a gray area. Adding a gray area will increase both sensitivity and specificity. Moreover, the presence of a gray area highlights the difficulties related to the early detection of cognitive impairment and mirrors the clinical reality quite accurately.

ACKNOWLEDGEMENTS

We gratefully acknowledge the help of the staff of the Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel, Switzerland. We thank Christian Brauchli, PhD, for methodological input and commenting on the manuscript, Michael Ehrensperger, PhD, and Marc Sollberger, MD, for providing advice during study conduction, and Ursi Kunze, MSc, for support in database management.

Conflicts of Interest and Source of Funding

Financial support for this research was provided from internal sources of the Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel, Switzerland and of the Department of Anesthesia, University Hospital Basel, Switzerland.

Authors' disclosures are available online. Dr. Monsch reports personal fees from AbbVie Pahrma, from Vifor Pharma, from Merz Pharma and from Roche Pharma outside the submitted work. Ms. Thomann reports personal fees from Vifor Pharma outside the submitted work.

Supplementary material

Supplementary table 1. Diagnoses in the patient sample.

Diagnoses	Mild NCD	Major NCD	Total
Alzheimer's clinical syndrome [1]	55	237	292
Vascular cognitive impairment [2, 3]	15	12	27
Lewy Body disease [4]	1	3	4
Behavioral variant fronto-temporal dementia [5]	0	4	4
Parkinson's disease [6]	3	5	8
Multiple system atrophy [7]	0	2	2
Progressive supranuclear palsy [8]	2	0	2
Primary progressive aphasia [9]	2	2	4
Posterior cortical atrophy [10]	0	2	2
Psychiatric disorder	7	1	8
Obstructive sleep apnea	4	0	4
Sleep disorder	3	0	3
Uncertain	39	9	48
Other	28	11	39

[1] Jack CR, Jr., Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2018;14:535-62.

[2] Sachdev P, Kalaria R, O'Brien J, Skoog I, Alladi S, Black SE, et al. Diagnostic criteria for vascular cognitive disorders: a VASCOG statement. Alzheimer Dis Assoc Disord. 2014;28:206-18.

[3] Smith E. Vascular Cognitive Impairment. Continuum (Minneapolis, Minn). 2016;22:490-509.

[4] McKeith IG, Boeve BF, Dickson DW, Halliday G, Taylor J-P, Weintraub D, et al. Diagnosis and management of dementia with Lewy bodies. Fourth consensus report of the DLB Consortium. 2017.

[5] Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, et al. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain. 2011;134:2456-77.

[6] Emre M, Aarsland D, Brown R, Burn DJ, Duyckaerts C, Mizuno Y, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. Mov Disord. 2007;22:1689-707; quiz 837.

[7] Palma JA, Norcliffe-Kaufmann L, Kaufmann H. Diagnosis of multiple system atrophy. Autonomic neuroscience : basic & clinical. 2018;211:15-25.

[8] Hoglinger GU, Respondek G, Stamelou M, Kurz C, Josephs KA, Lang AE, et al. Clinical diagnosis of progressive supranuclear palsy: The movement disorder society criteria. Mov Disord. 2017;32:853-64.

[9] Gorno-Tempini ML, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, et al. Classification of primary progressive aphasia and its variants. Neurology. 2011;76:1006-14.

[10] Crutch SJ, Schott JM, Rabinovici GD, Murray M, Snowden JS, van der Flier WM, et al. Consensus classification of posterior cortical atrophy. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2017;13:870-84.

Supplementary table 2. Empirically derived cut-offs and diagnostic accuracy estimates for the MoCA z-score.

Measure	**Mild+Major NCD vs. NF**	**Mild NCD vs. NF**
AUC (95% CI, DeLong)	**0.94 (0.91-0.96)**	**0.86 (0.81-0.91)**
Cut-□✉✲□◆♀ndex*	**-1.3**	**-1.1**
Correct classification rate†	88%	79%
Sensitivity (95% CI)	81% (80-82%)	68% (66-70%)
Specificity (95% CI)	94% (92-96%)	90% (88-92%)
Cut-off; 10th percentile in NCs	**-1.4**	**-1.4**
Correct classification rate†	88%	79%
Sensitivity (95% CI)	80% (79-81%)	61% (59-63%)
Specificity (95% CI)	96% (95-97%)	96% (95-97%)

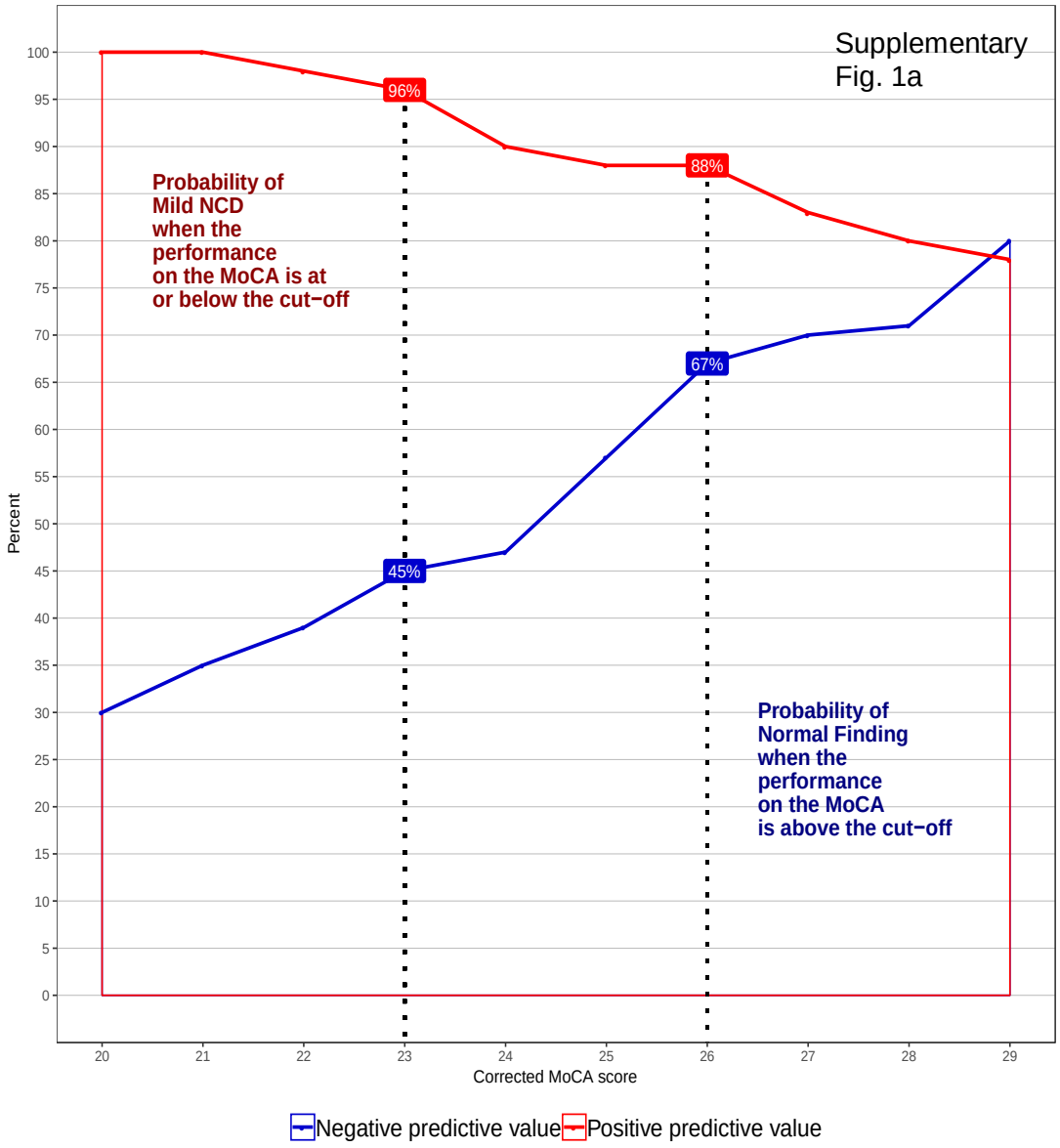
The MoCA z-score is a demographically corrected standard score based on the formula by Thomann, Goettel, Monsch et al. (2018).

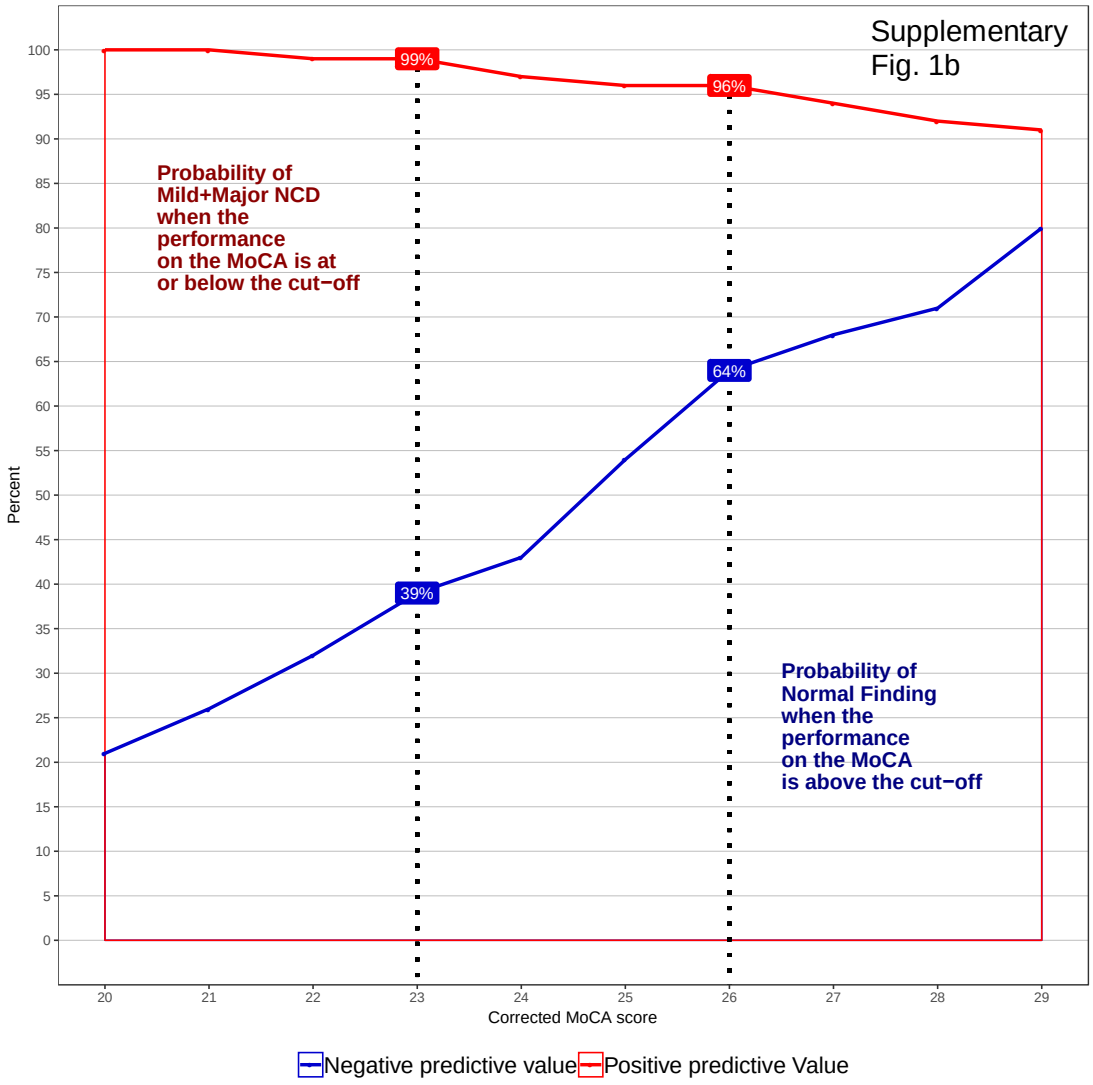
Abbreviations: AUC = Area under the curve; CI = Confidence interval; NC = Normal Controls; NCD = Neurocognitive disorder; NF = Normal Findings.

* Youden's index = Sensitivity + Specificity - 1

†Correct classification rate = (Sensitivity + Specificity)/2

Supplementary Figure 1. Positive and negative predictive values.





In Supplementary Fig. 1a, the positive predictive values (PPV) and negative predictive values (NPV) are plotted for Mild NCD vs. NF and highlighted for the proposed cut-offs of 23/24 points and 26/27 points. In Supplementary Fig. 1b, PPV and NPV are illustrated for Mild+Major NCD vs. NF. In all patient groups, PPV decrease and NPV increase with higher MoCA threshold scores. Again, using two separate cut-offs enhances both, PPV and NPV.

5. Study III:

Development of a Novel Self-administered Cognitive Assessment Tool and Normative Data for Older Adults

Raphael J. Monsch, MMed,[a] Amélie C. Burckhardt, MD,[b] Manfred Berres, PhD,[c] Alessandra E. Thomann, MSc,[a, d] Michael M. Ehrensperger, PhD,[d] Luzius A. Steiner, MD, PhD,[a, e] and Nicolai Goettel, MD[a, e]

Published

[a]Department of Anesthesia, Surgical Intensive Care, Prehospital Emergency Medicine and Pain Therapy, University Hospital Basel, University of Basel, Basel, Switzerland
[b]Department of Internal Medicine, Kantonsspital Baselland, Liestal, Switzerland
[c]Department of Mathematics and Technology, University of Applied Sciences Koblenz, Remagen, Germany
[d]Memory Clinic, University Center for Medicine of Aging Basel, Felix Platter Hospital, Basel, Switzerland
[e]Department of Clinical Research, University of Basel, Basel, Switzerland

R.J.M. and A.C.B. contributed equally.

The final publication is available at
https://journals.lww.com/jnsa/Abstract/2019/04000/Development_of_a_Novel_Self_administered_Cognitive.9.aspx

ABSTRACT

Background: Pre-existing cognitive impairment in surgical patients is one of the leading risk factors for adverse cognitive outcomes such as postoperative delirium and postoperative cognitive dysfunction. We developed a self-administered tablet computer application intended to assess the individual risk for adverse postoperative cognitive outcomes. This cross-sectional study aimed to establish normative data for the tool.

Methods: Healthy volunteers aged ≥ 65 years were administered the Mini-Mental State Examination (MMSE), Geriatric Depression Scale (GDS), and Consortium to Establish a Registry for Alzheimer's Disease-Neuropsychological Assessment Battery (CERAD-NAB) to assess cognitive health. All subjects completed the tablet computer application without assistance. Primary outcome measure was the test performance. Regression models were built for each cognitive domain score with the covariates age, gender, and education in cognitively healthy subjects. Demographically-adjusted standard scores (z-scores) were computed for each subtest.

Results: 283 participants (155 women, 128 men) were included in the final analysis. Participants' age was 73.8 ± 5.2 years (mean ± SD) and their level of education was 13.6 ± 2.9 years. MMSE score was 29.2 ± 0.9 points, GDS score was 0.4 ± 0.7 points, and CERAD-NAB total score was 98.7 ± 5.7 points. Older age was associated with poorer performance in the visual recognition task and in Trail Making Test B ($P < 0.05$ after Bonferroni-Holm adjustments).

Conclusions: This study provides normative data for a novel self-administered tablet computer application that is ultimately designed to measure the individual risk for adverse postoperative cognitive outcomes in elderly patients.

Trial Registration: ClinicalTrials.gov NCT02708823

Keywords: cognitive function; assessment; postoperative delirium; postoperative cognitive dysfunction; tablet computer application; normative data

5.1 INTRODUCTION

In light of a growing geriatric patient population, health care professionals are increasingly faced with specific challenges of elderly patients in the primary care and hospital setting. The need for surgical procedures increases with patient age (Hall et al., 2010). Elderly patients undergoing surgery are more vulnerable to adverse postoperative outcomes due to advanced age, frailty, and concomitant medical conditions (Story et al., 2010). Adverse cognitive outcomes such as postoperative delirium (POD) and postoperative cognitive dysfunction (POCD) are frequently encountered in older surgical patients and are associated with increased morbidity and mortality (Sanders et al., 2011; Steinmetz et al., 2009; Witlox et al., 2010). An early identification of risk factors is useful for the targeted prevention of cognitive disorders in hospitalized patients (Inouye et al., 1999). While most predictors for POD and POCD may be detected in the medical history, clinical examination, or laboratory investigations, some may be missed in the absence of a specific assessment. Pre-existing cognitive impairment in surgical patients is one of the strongest risk factors for further postoperative cognitive decline including POD (Dasgupta & Dumbrell, 2006; Inouye et al., 2014; Jones et al., 2006; Sprung et al., 2017) and POCD (Nadelson et al., 2014; Silbert et al., 2015). However, it tends to be underdiagnosed (Prince, Bryce, & Ferri, 2011; Young, Meagher, & Maclullich, 2011), because an objective evaluation of the cognitive performance is time-consuming and usually requires trained personnel. Therefore, it may be challenging to implement the routine assessment of cognitive status in all geriatric patients presenting for surgery (C. Brown & Deiner, 2016). Besides, most cognitive screening tools available to date are not specifically intended for preoperative use in surgical patients (Long, Shapiro, & Leung, 2012). Some current risk prediction models for POD do not include the assessment of cognitive functions at baseline (Evered, 2017; Lee et al., 2017).

Our goal was to create a new tool to assess individual baseline cognition as a major risk factor for adverse postoperative cognitive outcomes in surgical patients. Key requirements for the design of the CogCheck application were self-administration, user-friendliness, language-free content (pictures), conciseness (i.e., administration time < 30 minutes), and automated scoring. These may facilitate routine use in clinical practice (e.g., during preoperative evaluation for anesthesia) and offer potential advantages over other screening tools. Eventually, the purpose of the CogCheck application is to simplify and standardize preoperative cognitive

testing in the elderly. Compared to CogCheck, other preoperative cognitive assessments do not use computerized testing (Long et al., 2012), which may be beneficial regarding test reliability and scoring. In addition, the self-administrative character of CogCheck and the possibility of remote and parallel testing may reduce personnel and resource costs.

The development of such tool involves several steps: (1) identification of relevant cognitive domains, (2) choice of task to assess these domains, (3) computer programming of the tasks, (4) pilot study to assess applicability of the tool, and (5) collection of normative data in a group of individuals with established cognitive health (Crook, Kay, & Larrabee, 2009). Once these steps have been carried-out successfully, the new tool may be used in a series of validation studies. The objective of this cross-sectional study was to collect normative data in cognitively healthy individuals, and find the adjustment necessary to eliminate the influence of demographic characteristics (age, gender, and education).

5.2 MATERIALS AND METHODS

5.2.1 Study design

We conducted a cross-sectional study to acquire normative data for the tablet computer-based application, CogCheck. Ethical approval for this study (protocol N° EKNZ BASEC 2016-00393) was provided by the institutional ethics board (*Ethikkommission Nordwest- und Zentralschweiz*) on April 12, 2016. A substantial amendment to the study protocol was approved on November 11, 2016. All study participants provided written informed consent. The study was conducted in respect of the most recent version of the Declaration of Helsinki and registered on ClinicalTrials.gov (NCT02708823) prior to data acquisition. This manuscript adheres to the applicable EQUATOR network guidelines.

5.2.2 Participants and setting

All study participants were healthy nonsurgical volunteers recruited from the Registry of Individuals Interested to Participate in Research established by the Memory Clinic, University Center for Medicine of Aging Basel, Felix Platter Hospital, in Basel, Switzerland. Only subjects who had previously filled out a

standardized medical questionnaire were considered. Data from eligible participants were screened for inclusion and exclusion criteria. Inclusion criteria were: (1) age ≥ 65 years, (2) education ≥ 7 years, (3) fluency in the German language, and (4) written informed consent. Exclusion criteria were: (1) history of cognitive impairment, (2) signs of depression, (3) severe sensory or motor impairment interfering with cognitive testing, (4) serious somatic disease, disease or event affecting the central nervous system (head trauma with loss of consciousness > 5 minutes, any brain surgery, general anesthesia within the last 3 months, alcoholism, intoxication with neurotoxic substances), (5) cerebrovascular disease, (6) regular medication with psychoactive drugs except for benzodiazepines, and (7) participation in any cognitive study within the last 3 months or previous participation in a study using CogCheck.

In order to ensure cognitive health of participants, only those with at least 27/30 points (Thalmann et al., 2002) in the Mini-Mental State Examination (MMSE; Folstein et al., 1975) and more than 85.89 points (Ehrensperger et al., 2010) in the German version of the Consortium to Establish a Registry for Alzheimer's Disease-Neuropsychological Assessment Battery (CERAD-NAB; Morris et al., 1989) were included. Subjects with more than 5/15 points on the brief version of the Geriatric Depression Scale (GDS; Pocklington, Gilbody, Manea, & McMillan, 2016), indicating signs of depression, were excluded. Optimal homogeneity of the study population was achieved by stratification of participants according to age and gender categories.

5.2.3 Design of CogCheck

The CogCheck application was developed in a joint project by the Department of Anesthesia at University Hospital Basel and the Memory Clinic at Felix Platter Hospital in Basel, Switzerland. Since objective assessment of a patient's cognitive status is highly resource-dependent (Saxton et al., 2009), our goal was to create a computerized risk-stratification tool for adverse postoperative cognitive outcomes in surgical patients that is easy to use and does not require trained personnel. Previous investigations showed that even persons without computer experience were able to perform well using computer-based tests (Fazeli, Ross, Vance, & Ball, 2013). Moreover, study subjects were more successful when using a tablet computer with touch screen instead of a computer with a mouse or a keyboard

(Saxton et al., 2009; Werner, Werner, & Oberzaucher, 2012). Thus, we designed a tablet computer application in which all subtests are language-free. Instructions – which can be easily translated into other languages – are provided in writing and are complemented with short videos. This also allows for the assessment of patients with hearing impairment.

We compared existing preoperative risk scores (Freter et al., 2005; S. K. Inouye, Viscoli, Horwitz, Hurst, & Tinetti, 1993; Marcantonio et al., 1994) to decide which predictors should be included in our new tool. The final version of CogCheck (see Figure, Supplemental Digital Content 1, which shows translated screenshots of the application) used for test standardization included: (1) demographic and medical data (sensory impairment [Kalisvaart et al., 2006), age [National Clinical Guideline (NGC), 2010; Wimo et al., 2017], medications [Goldenberg et al., 2006], education [Jones et al., 2006]), (2) cognitive self-assessment (NGC, 2010), (3) temporal orientation (Folstein et al., 1975; Long et al., 2012), and (4) a set of 7 automated subtests of cognitive functions (visual recognition [Benton, 1972], picture learning and recognition [Saxton et al., 2009], digit span [Erlanger et al., 2002], spatial span [Brunetti, Del Gatto, & Delogu, 2014], reaction time and attention [Saxton et al., 2009], and Trail Making Tests [TMT] A and B [Salthouse & Fristoe, 1995]). The automated scoring system for CogCheck is based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, 2013) definitions for neurocognitive disorders.

At an initial stage, user-friendliness of CogCheck was evaluated in a pilot project (Burckhardt, 2014) with 20 cognitively healthy volunteers (10 women, 10 men; mean age: 71.8 ± 3.4 years; mean Montreal Cognitive Assessment [MoCA; Nasreddine et al., 2005] score: 28.0 ± 0.9 points) and 13 cognitively impaired patients (5 women, 8 men; mean age: 76.5 ± 4.5 years; mean MoCA score: 22.3 ± 2.6 points) of the Memory Clinic. Twenty-seven pilot study participants (82%) privately owned and used a computer, 7 (21%) a tablet computer, and 9 (27%) a smartphone. The majority of cognitively healthy and impaired subjects were able to successfully complete the assessment without or with minimal help (95% and 85%, respectively). The CogCheck application received high overall quality and acceptance ratings (clear layout: 97%; easy navigation: 88%). Successively, some practical features of the application were improved (e.g., font size, color coding, and touchscreen sensibility).

A second pilot study (Anyiam, 2018) examined CogCheck in older surgical patients and evaluated the applicability of the tool in a real-life clinical setting. Forty-six

patients (29 women, 17 men; mean age: 73.3 ± 5.6 years; mean MMSE score: 28.3 ± 1.2 points) scheduled for major surgery completed the CogCheck application in our anesthesia preoperative evaluation clinic. All patients were able to complete testing without assistance. During the first five days after surgery, patients were assessed for POD using the Delirium Rating Scale-Revised-98 (DRS-R-98; Trzepacz et al., 2001). When applying the formal DRS-R-98 cut-off scores, no patient was found to be affected by POD in this cohort (mean maximum DRS-R-98 score: 5.4 ± 2.9 points; range: 0–12 points). This was possibly due to sampling bias. Consequently, this data set was insufficient to validate CogCheck in surgical patients.

5.2.4 Variables and data sources

Study participants were examined by one of four individually trained psychology master's students in a quiet room, seated at a table. After obtaining consent, the examiner first updated the individual's medical questionnaire and medication list. Second, the MMSE and GDS were administered. Subjects then performed CogCheck on an iPad Air tablet computer with 9.7-inch display using iOS 10.2 or 10.3 (Apple Inc., Cupertino, CA, USA). Although the examiner remained in the room during CogCheck testing, he or she was not allowed to interact in any way with the subject. Finally, the extended German version of the CERAD-NAB (Schmid, Ehrensperger, Berres, Beck, & Monsch, 2014) was administered.

Data from CogCheck were sent in real-time to a secure server at University Hospital Basel using a locked Wireless Local Area Network (WLAN) connection. Examiners were blinded to application data. Paper-based study data were recorded directly onto the case report form and later transferred into an electronic database using FileMaker Pro (FileMaker Inc., Santa Clara, CA, USA).

5.2.5 Statistical analysis

We evaluated the effects of common demographic characteristics on test performance and examined the distribution of scores. First, 20 regression models for each cognitive subtest (see Table, Supplemental Digital Content 2, which displays the content and structure of the CogCheck application) were calculated with the covariates age, gender, education, their interactions, and their potential

nonlinear relationships using quadratic terms (Berres et al., 2008). The optimal model was determined by leave-one-out cross-validation, i.e., minimizing the Prediction Residual Sums of Squares (PRESS) statistics among the 20 regression models for each response variable (Berres et al., 2008). Second, if necessary, optimal transformations (Box-Cox family or arcsine) were applied to achieve normality and homoscedasticity of the residuals. Third, step one was repeated with transformed variables determining an optimal model from the 20 models. This was always the same or a similar model as in step one, which speaks for a certain robustness of the analysis. Finally, formulae for demographically-adjusted standard scores (z-scores) were computed based on the final regression model. The Bonferroni-Holm method for multiple testing was applied in order to estimate the hypothetical effects of age and education in all subtests.

In order to estimate the 5^{th} and 95^{th} percentile with a maximum deviation of 2% for the normative data (Jennen-Steinmetz & Wellek, 2005), at least 171 subjects were needed. Age, gender, and education were predefined as predictor variables, and three additional predictor variables with interactions and quadratic terms were anticipated. Ten subjects per predictor variable were included to account for adjustments in the regression models. Hence, the minimum sample size was 231. All statistical analyses were performed using R, version 3.4.1 (R Foundation, Vienna, Austria).

5.3 RESULTS

5.3.1 Participants

All study-related examinations took place between December 2016 and April 2017. At the time of the study, the Registry of Individuals Interested to Participate in Research counted 2162 volunteers, including 794 subjects who had filled out a standardized medical questionnaire. Of 487 eligible subjects who were contacted by letter, 334 were included in the study. The final sample for analysis consisted of 283 cognitively healthy volunteers (155 women, 128 men). Figure 1 shows the process of recruitment and inclusion in detail. For the final sample (n = 283), mean subject age was 73.8 ± 5.2 (range 65–91) years, and mean education was 13.6 ± 2.9 (range 7–20) years. Each age category was represented by at least 21 subjects per gender. The study population comprised nearly equal numbers of men and

women in each age category. Demographic characteristics, medical comorbidities, and neuropsychological test results of participants are summarized in Table 1. There was no missing data on the key variables in our main analysis.

Table 1. Demographic characteristics according to age category

	All participants (n = 283)	**65–69 years (n = 68)**	**70–74 years (n = 102)**	**75–79 years (n = 68)**	**> 79 years (n = 45)**
Age; years	73.8 (5.2)	67.6 (1.4)	72.2 (1.3)	76.5 (1.4)	82.6 (2.4)
Male gender; n (%)	128 (45.2)	26 (38.2)	44 (43.1)	34 (50.0)	24 (53.3)
Education; years	13.6 (2.9)	13.2 (2.7)	14.0 (2.8)	13.7 (3.1)	13.3 (2.8)
MMSE score; points	29.2 (0.9)	29.4 (0.7)	29.3 (0.9)	29.0 (0.9)	28.9 (1.0)
GDS score; points	0.4 (0.7)	0.3 (0.8)	0.4 (0.7)	0.3 (0.6)	0.4 (0.7)
CERAD-NAB-Plus total score; points	98.7 (5.7)	97.9 (5.5)	98.6 (5.2)	99.5 (5.9)	99.0 (6.5)

Data are presented as mean (SD), unless otherwise stated. CERAD-NAB = Consortium to Establish a Registry for Alzheimer's Disease-Neuropsychological Assessment Battery; GDS = Geriatric Depression Scale; MMSE = Mini-Mental State Examination.

Fig. 1. Recruitment process.

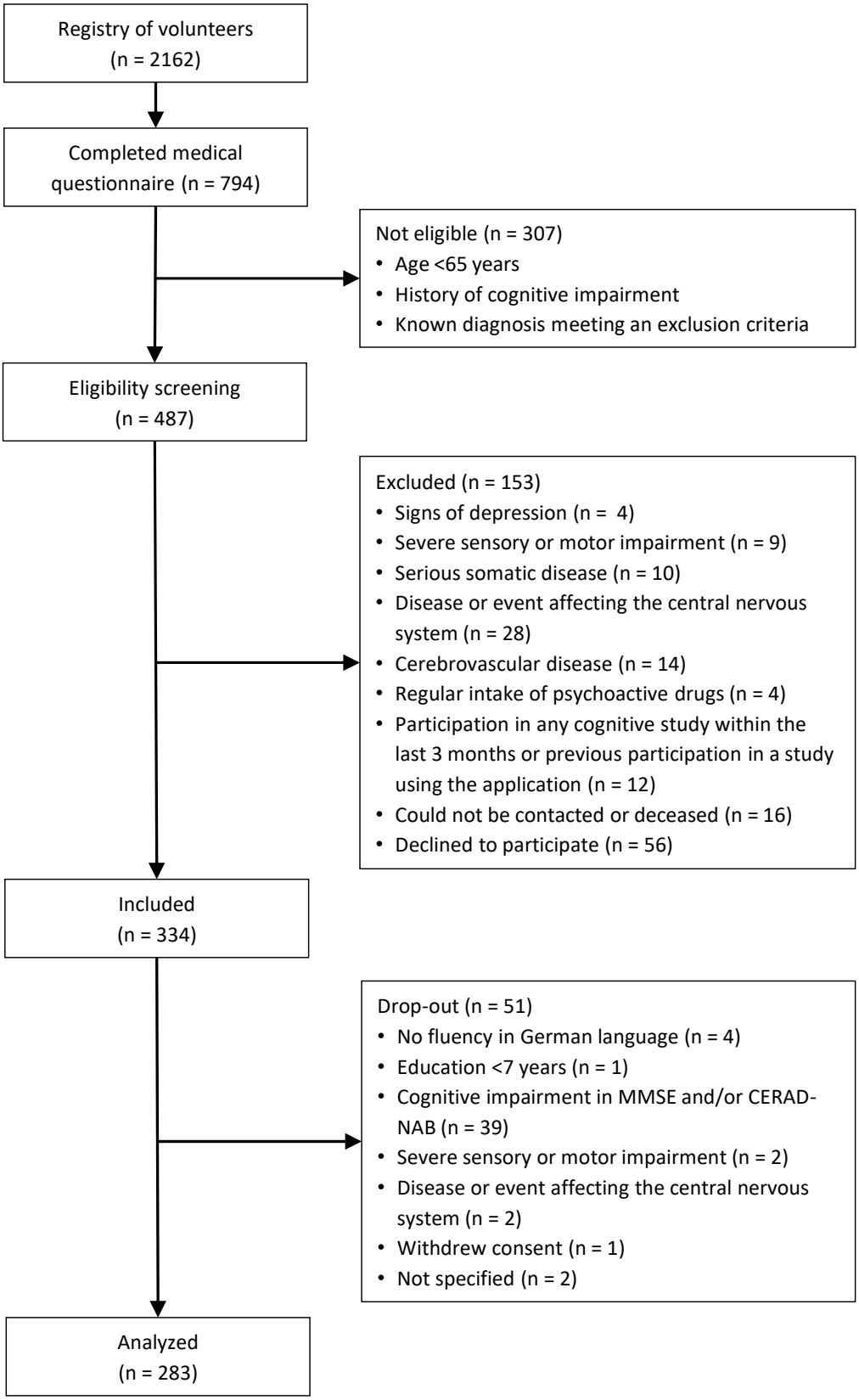


Table 2. CogCheck test results

	All participants (n = 283)	65–69 years (n = 68)	70–74 years (n = 102)	75–79 years (n = 68)	> 79 years (n = 45)
Demographic and medical data; n (%)					
Sensory impairment					
Use of vision aids	272 (96.1)	66 (97.0)	95 (93.1)	66 (97.1)	45 (100)
Presence of hearing impairment	121 (42.8)	22 (32.4)	34 (33.3)	32 (47.1)	33 (73.3)
Daily drug intake					
No drugs	60 (21.2)	18 (26.5)	25 (24.5)	11 (16.2)	6 (13.3)
1 to 3 drugs	172 (60.8)	46 (67.7)	62 (60.8)	40 (58.8)	24 (53.3)
4 to 7 drugs	42 (14.8)	4 (5.9)	12 (11.8)	15 (22.1)	11 (24.4)
>7 drugs	9 (3.2)	0 (0)	3 (2.9)	2 (2.9)	4 (8.9)
Age entered correctly[a]	245 (86.6)	60 (88.2)	92 (90.2)	60 (88.2)	33 (73.3)
Education entered correctly	137 (48.4)	33 (48.5)	48 (47.1)	31 (45.6)	25 (55.6)
Language					
Native German speaker	277 (97.9)	68 (100)	100 (98.0)	66 (97.1)	43 (95.6)
Other, but fluent in German	6 (2.1)	0 (0)	2 (2.0)	2 (2.9)	2 (4.4)
Cognitive self-assessment[b]					
Memorizing new things	2.6 (0.5)	2.7 (0.5)	2.6 (0.5)	2.6 (0.6)	2.4 (0.5)
Remembering names	2.4 (0.6)	2.5 (0.6)	2.4 (0.6)	2.4 (0.6)	2.3 (0.6)
Multiple simultaneous tasks	2.8 (0.4)	2.9 (0.4)	2.8 (0.5)	2.7 (0.5)	2.7 (0.4)
Financial issues	3.1 (0.3)	3.0 (0.2)	3.1 (0.4)	3.0 (0.3)	3.0 (0.4)
Remembering appointments	3.0 (0.3)	3.0 (0.3)	3.0 (0.3)	3.0 (0.2)	2.9 (0.3)
Temporal orientation;[c] n (%)					
Weekday entered correctly	281 (99.3)	68 (100)	101 (99.0)	67 (98.5)	45 (100)
Date entered correctly	231 (81.6)	56 (82.4)	81 (79.4)	58 (85.3)	36 (80.0)
Automated subtests of cognitive functions					
Visual recognition; raw score[d]	12.0 (1.9)	12.3 (1.7)	12.2 (1.7)	12.0 (1.7)	10.8 (2.3)
Picture recognition; raw score[d]	27.5 (2.1)	28.1 (1.8)	27.5 (2.1)	26.7 (2.3)	27.5 (2.0)
Spatial span; raw score[d]	7.0 (1.7)	7.2 (1.7)	7.1 (1.7)	7.2 (1.6)	6.2 (1.8)
Digit span; raw score[d]	8.4 (2.0)	8.7 (2.1)	8.4 (2.1)	8.6 (1.9)	7.5 (1.8)
TMT-A; number of line connections/min	21.9 (5.9)	23.4 (6.1)	22.2 (6.2)	21.7 (5.6)	19.3 (3.7)
TMT-B; number of line connections/min	15.1 (4.1)	16.5 (3.3)	15.7 (3.9)	14.8 (4.2)	12.3 (3.9)

Data are presented as mean (SD), unless stated otherwise. TMT = Trail Making Test.

[a]Error analysis showed that 97.4% of subjects, who had entered an incorrect age, had rounded their age up to the next year.

[b]Cognitive functions were self-assessed on a five-point Likert scale (1 = much worse, 2 = somewhat worse, 3 = no change, 4 = somewhat better, 5 = much better) compared to two years ago.

[c]Error analysis showed that 0.7% of subjects entered the weekday incorrectly, 17.6% entered the day incorrectly, none entered the month incorrectly, and 1.1% entered the year incorrectly.

[d]Possible range of values is 0 to 15 for visual recognition, 0 to 30 for picture recognition, 0 to 16 for spatial span, and 0 to 18 for digit span.

5.3.2 Results of CogCheck

Demographic data, self-assessment of cognitive functions and testing of temporal orientation originating from CogCheck are summarized in Table 2. The mean time necessary to complete the application was 21.7 ± 2.2 minutes. All participants were able to successfully complete the assessment without help.

The influence of age, gender, and education on other subtests was not uniformly negative or positive due to interactions and quadratic effects. When applying the Bonferroni-Holm adjustment to test the effect of age and education in six subtests ($\alpha = 0.05$), the age effect was significant in the visual recognition task and TMT-B. The education effect was significant in the visual recognition task and fell just short of significance in TMT-B (adjusted $P = 0.054$). Because of modifying effects (interactions and quadratic terms), uniform effects of age and education could not be tested in the other four subtests.

5.3.3 Calculation of standard scores

For each cognitive subtest of CogCheck, we chose the best predictive model and computed demographically-adjusted standard scores (z-scores). The basic formula for the calculation of standard scores (z = [transformed score – expected score]/ residual standard error) was applied for each cognitive subtest (Table 3). The Figure in Supplemental Digital Content 3 provides the detailed analysis of all CogCheck subtests.

Table 3. Formulae for demographically-adjusted standard scores

Cognitive subtest	Standard score formula
Visual recognition	$z = ((RS - 2)^{1.5} - (54.694 - 0.398 \times A + 0.479 \times E)) / 8.308$
Picture recognition	$z = (\text{asin}(\text{sqrt}(RS / 30.5)) - (1.460 - 0.0042 \times A + 0.0055 \times E + 0.056 \times G + 0.0006 \times (A - A_{mean}) \times G)) / 0.115$
Spatial span	$z = (RS^{1.4} - (31.811 - 0.245 \times A + 0.141 \times E - 0.964 \times G + 0.185 \times (A - A_{mean}) \times G + 0.050 \times (E - E_{mean})^2)) / 5.031$
Digit span*	$z = (RS - (11.736 - 0.053 \times A + 0.087 \times E - 0.811 \times G - 0.0079 \times (A - A_{mean})^2)) / 1.923$
TMT-A	$z = (RS^{0.75} - (18.48 - 0.0912 \times A - 0.136 \times E + 0.064 \times G + 0.241 \times (E - E_{mean}) \times G + 0.026 \times (E - E_{mean})^2)) / 1.998$
TMT-B	$z = (RS^{1.5} - (153.17 - 1.488 \times A + 1.2636 \times E)) / 21.653$

The basic formula for the calculation of standard scores is z = (transformed score - expected score) / residual standard error.

*No transformation was necessary to receive normal distribution for the digit span score.

A = age; E = education; G = gender; RS = raw score; TMT = Trail Making Test.

5.4 DISCUSSION

The CogCheck application is a completely self-administered cognitive assessment and screening tool intended for use in surgical patients. This cross-sectional study provides demographically-adjusted normative data for the CogCheck tool. Taking into account age, gender, and education, we calculated standard scores for six cognitive subtests that, in combination, may provide an indication of the overall cognitive status. Two previous pilot studies independently demonstrated the user-friendliness and applicability of CogCheck in cognitively healthy and impaired subjects, as well as in surgical patients.

Pre-existing cognitive impairment is reported to have a significant impact on the incidence of adverse postoperative cognitive outcomes. In earlier studies, the odds ratio for delirium ranged from 6.3 (95% confidence interval [CI] 2.9–13.7) up to 11.5 (95% CI 6.1–20.1) in patients suffering from cognitive impairment (Wimo et al., 2017). For POCD, the odds ratio was 2.4 (95% CI 1.1–5.5) (Silbert et al., 2015). After validation of CogCheck in surgical patients, the tool may eventually screen for cognitive impairment as a major risk factor for adverse postoperative cognitive outcomes via self-administered testing on a tablet computer.

Participants of the current normative study were cognitively healthy volunteers, and relatively strict exclusion criteria (cut-off scores for MMSE, CERAD-NAB, and GDS)

were applied. This eliminates potential confounders (presence of mild cognitive impairment, dementia, or depression) and leads to almost ideal normative data. Hence, clinicians may better interpret test results of patients affected by conditions associated with poor cognitive performance. Subjects with medical comorbidities commonly found in the elderly (Table 1) were intentionally not excluded from the study for a better representation of the geriatric population.

We used a regression-based analysis to calculate normative data for each subtest of the assessment application. This approach considers specific demographical data that are critical for an appropriate estimation of the individual performance and does not rely on categories (e.g., age groups) which are somewhat arbitrary. This increases the diagnostic accuracy in subjects at the extremes of such groups.

The composition of different cognitive tests in CogCheck may result in a more adequate assessment, as cognitive impairment and dementia may affect different domains of cognition. Assessing a smaller number of domains for the benefit of time may not capture the complete picture of cognitive impairment. The CogCheck application, in turn, has a multidimensional character.

Limitations of our study include the potential selection bias resulting in super-optimal normative data. Participants included in this study were recruited from an existing registry of nonsurgical volunteers. These individuals might have a higher intellect or display a greater motivation to perform well in cognitive testing than the average population. This bears the risk of overestimating cognitive impairment if interpretation of individual performance is missed. Therefore, our normative data must be considered as a guideline, and test results of patients from very different cultural backgrounds or individuals with very low education require cautious interpretation. It was decisive to include only individuals with established cognitive health, since their scores serve as starting points for the interpretation of scores from actual patients. This healthy normative sample will not be representative of older adults requiring surgery in all aspects, and expected differences will have to be explained on clinical grounds. Finally, since CogCheck was envisioned as a one-time screening test, we did not study the possibility of repeated/longitudinal assessment and tracing of a perioperative cognitive trajectory in individual patients, as well as the test-retest reliability.

Normative data are essential for any assessment tool, even when a traditional examiner-administered test is programmed for use on a computer, as it becomes a

new and different test (Bauer et al., 2012). The general question arises whether computerized assessment is appropriate for use in the elderly. One could assume that elderly people who are used to electronic devices may achieve better test results than those who are not, or do not feel comfortable using computers. However, previous findings suggest that the level of computer experience among older adults is not associated with the performance in a computerized test (Fazeli et al., 2013). Moreover, a recent literature review showed that people with dementia are able to independently use touchscreen technology (Joddrell & Astell, 2016). A disadvantage of computerized testing is the absence of an opportunity to motivate the patient as an examiner would be able to do. Nevertheless, self-administration is more resource-efficient and eliminates potential rater-related bias.

Traditional neuropsychological assessment batteries such as the CERAD-NAB are strongly based on verbal language. In contrast, the cognitive subtests in CogCheck are entirely language-free. A number of automated tools to assess cognitive functions also require the presence of a bedside examiner, include tests with a computer-generated voice (which can be difficult for patients with impaired hearing), or need handling of hardware (stylus, computer mouse, or keyboard). Some high-quality computerized applications (Zygouris & Tsolaki, 2015) like COGNIGRAM (CogState Ltd), CANTAB Mobile (Cambridge Cognition Ltd; Robbins et al. [1994]), or the NIH Toolbox (Health Measures) require the purchase of a license. However, considering recent health care resource cuts, paid single assessments may hinder the broad use of these tools in clinical practice (Weir et al., 2014; Zygouris & Tsolaki, 2015). We plan to make the CogCheck application available for free to any interested clinician and researcher. While some assessment tools take longer, the average time of 21.7 minutes needed to complete CogCheck seems reasonable. In addition, our tool screens for preoperative risk factors beyond pre-existing cognitive impairment (e.g., polymedication).

The current European and American guidelines on adverse postoperative cognitive outcomes recommend preoperative screening for risk factors including mental status for any patient without known history of cognitive impairment (Aldecoa et al., 2017; Mohanty et al., 2016). Preoperative screening may not only help to identify vulnerable patients but also guide preventive strategies. A standardized cognitive evaluation before surgery may offer important baseline information in patients experiencing postoperative cognitive decline. Still, the implementation of routine

screening for cognitive impairment in surgical patients may be challenging in daily practice (van Meenen, van Meenen, de Rooij, & ter Riet, 2014).

Since this study first provides normative data for an elderly nonsurgical population, CogCheck application data may not yet be used in a risk prediction model of adverse postoperative cognitive outcomes in surgical patients. Validation of the CogCheck application with postoperative outcome data is necessary before it may fully enter clinical routine. We plan to investigate the association of CogCheck performance and POD incidence in a follow-up study of patients undergoing cardiac surgery. Succeeding validation of the tool, CogCheck opens a wide field of research. Identifying vulnerable patient populations may simplify the study of targeted preventive measures to reduce the incidence of adverse postoperative cognitive outcomes (e.g., nonpharmacological multicomponent strategies [Inouye et al., 1999], cognitive or physical prehabilitation, prophylactic medication, and perioperative anesthetic considerations). Automatic data integration with digital records (e.g., medical history, medication lists, laboratory values, type of surgery) is conceivable in the future.

5.5 CONCLUSIONS

This study in healthy nonsurgical volunteers provides normative data for the CogCheck cognitive assessment tool. The CogCheck application measures the individual cognitive performance adjusted for demographic influences. However, clinical implementation of CogCheck to identify surgical patients with a high risk for adverse postoperative cognitive outcomes will only be possible after validation of the tool. In future research directed at the targeted prevention of adverse postoperative cognitive outcomes, this simple self-administered assessment tool may provide important information regarding the preoperative cognitive status.

ACKNOWLEDGEMENTS

We gratefully acknowledge the help of all participating volunteers, as well as the staff of the Department of Anesthesia, Surgical Intensive Care, Prehospital Emergency Medicine and Pain Therapy, University Hospital Basel, and the staff of the Memory Clinic, University Center for Medicine of Aging Basel, Felix Platter Hospital, Basel, Switzerland. In particular, we thank Panagiota Mistridis, PhD, and Sabine Krumm, PhD, for providing advice during study conduction, and Ursi Kunze, MSc, for her support in database management. We thank Andrina Baitella, BSc, Sarah Roesiger, BSc, Jael Fasnacht, BSc, and Antonino Cusano, BSc, for their help in conducting the study interviews. We thank Simon Wunderlin for information technology support and Allison Dwileski, BSc, for proof-reading the manuscript.

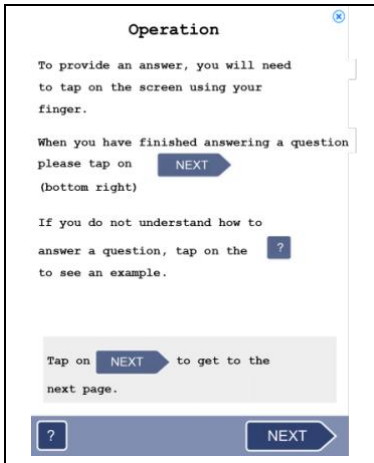
This report describes a cross-sectional study. The authors state that the report includes the items in the STROBE checklist for cross-sectional studies. The manuscript was screened for plagiarism with iThenticate.

Conflicts of Interest and Source of Funding

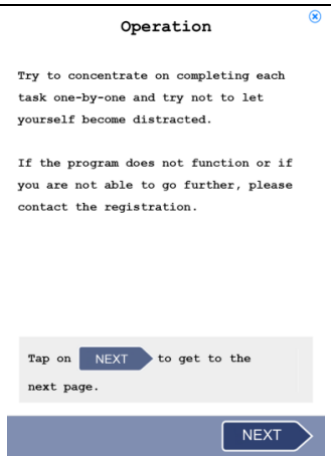
This research was supported by departmental sources. L.A.S. has received speaker honoraria from Medtronic Schweiz (Münchenbuchsee, Switzerland), Covidien (Neuhausen am Rheinfall, Switzerland) MSD, (Luzern, Switzerland), Hamilton Medical (Bonaduz, Switzerland), Lilly (Vernier, Switzerland), and Orion Pharma (Zug, Switzerland). None of these fees were related to this study. N.G. is currently receiving a grant (pp-16-01) from the Propatient Research Foundation (Basel, Switzerland). The remaining authors have no conflicts of interest to disclose.

Supplementary material

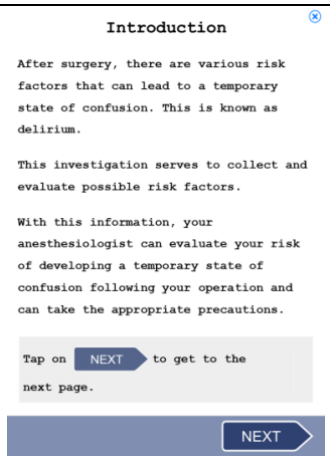
Supplemental digital content 1. Screenshots of CogCheck.



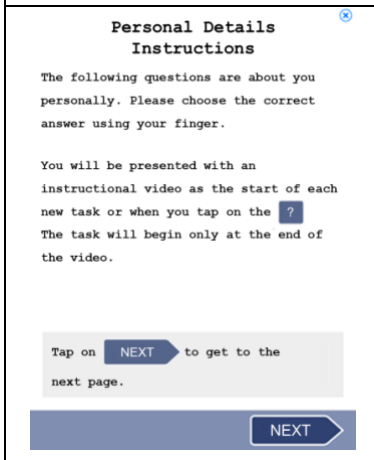
Introduction



Introduction

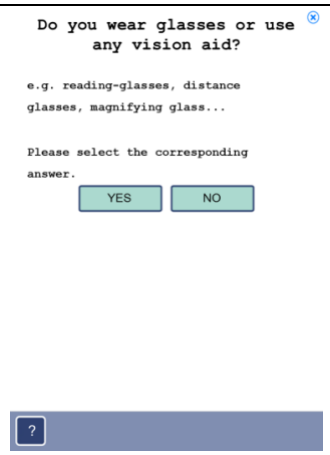


Introduction

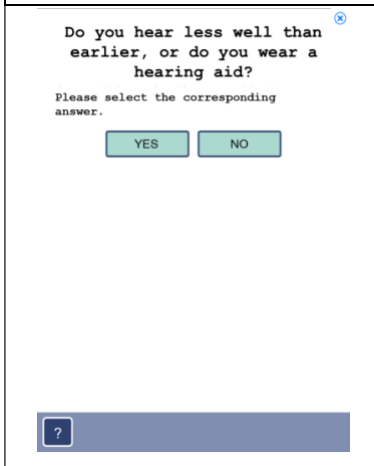


Instructions Part 1

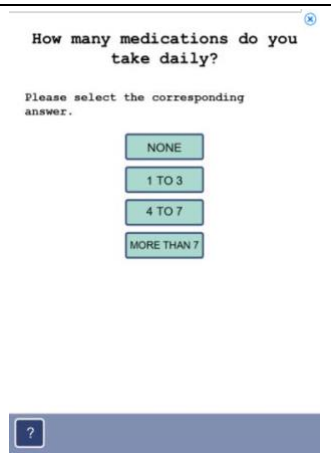
Instruction video



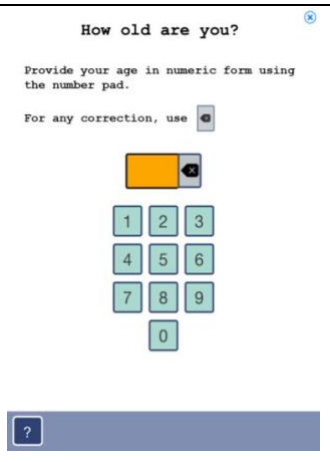
Vision



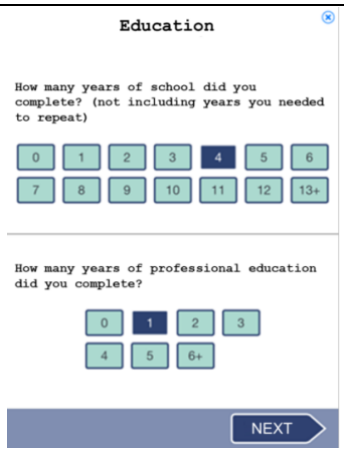
Hearing



Medication

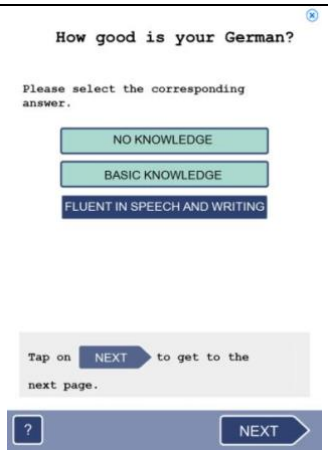


Age

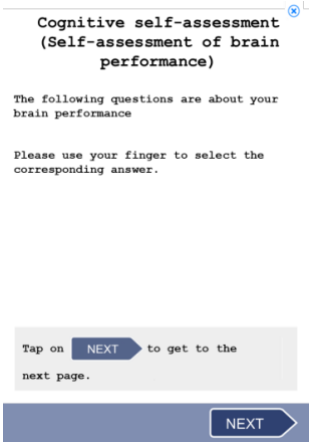


Education

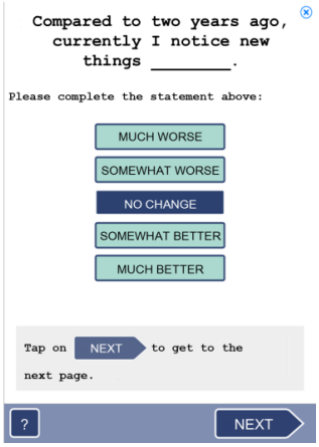
Language



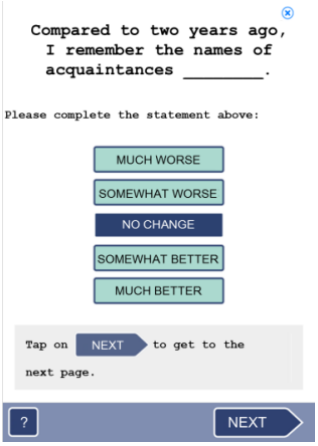
German language skills



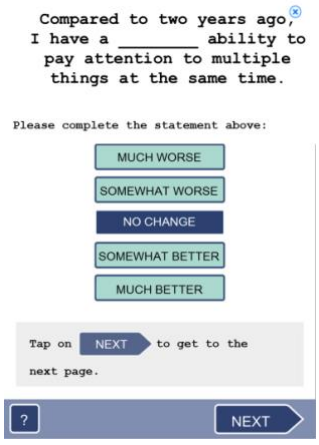
Cognitive self-assessment



Cognitive self-assessment:
Memorizing new things



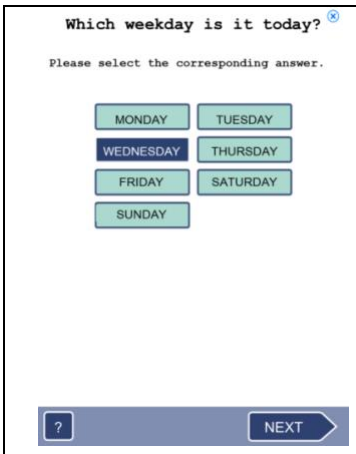
Cognitive self-assessment:
Remembering names



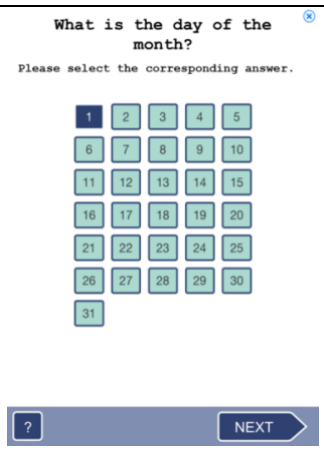
Cognitive self-assessment:
Multiple simultaneous tasks

Cognitive self-assessment:
Financial issues

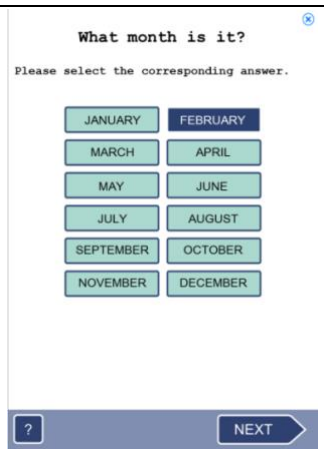
Cognitive self-assessment:
Remembering appointments



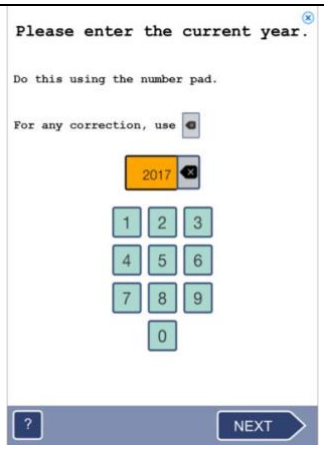
Temporal orientation: Weekday



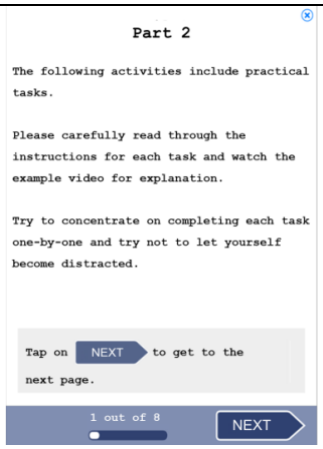
Temporal orientation: Day



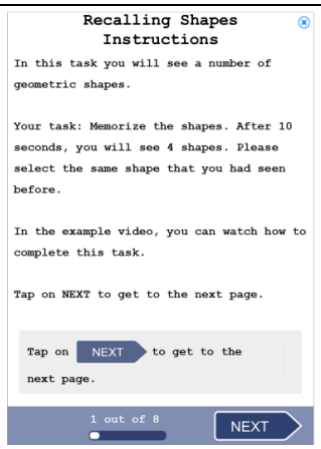
Temporal orientation: Month



Temporal orientation: Year

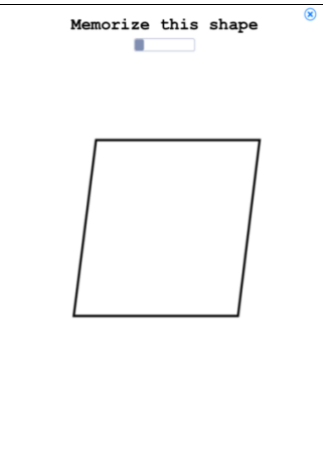


Instructions Part 2

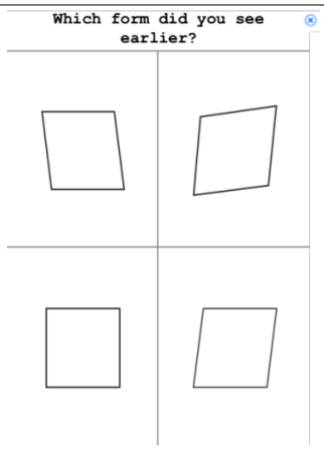


Visual recognition

Visual recognition: Video



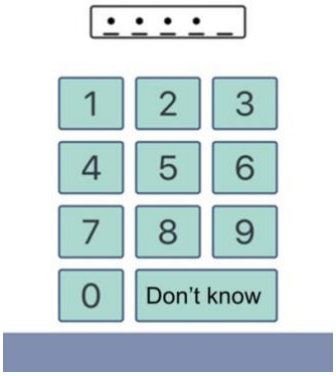
Visual recognition: Presentation



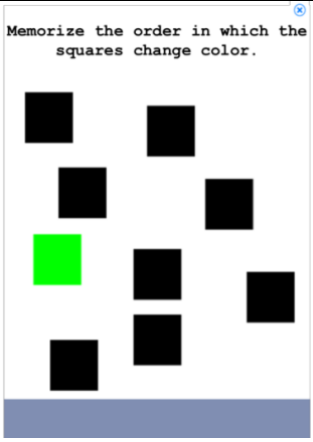
Visual recognition: Selection

Picture learning and recognition:
Memorizing images

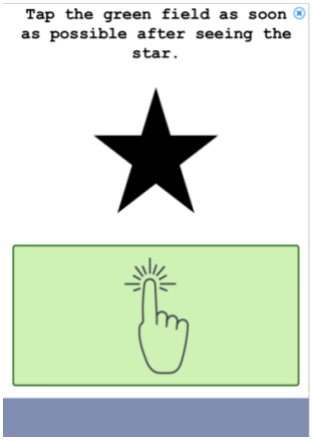
Digit span:
Memorizing numerical series



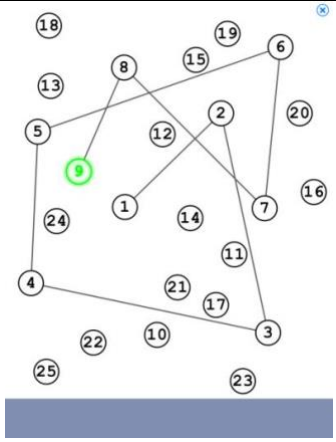
Digit span:
Recalling numerical series



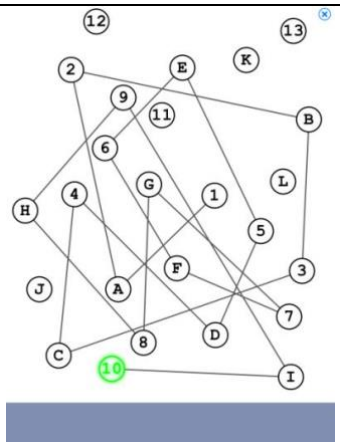
Spatial span



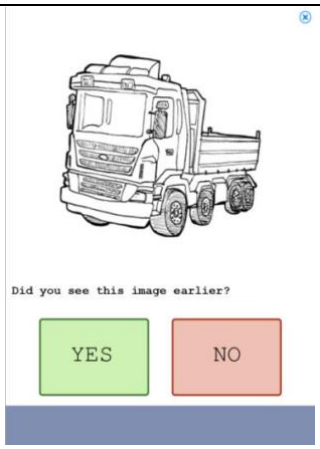
Reaction time and attention



Trail Making Test A



Trail Making Test B

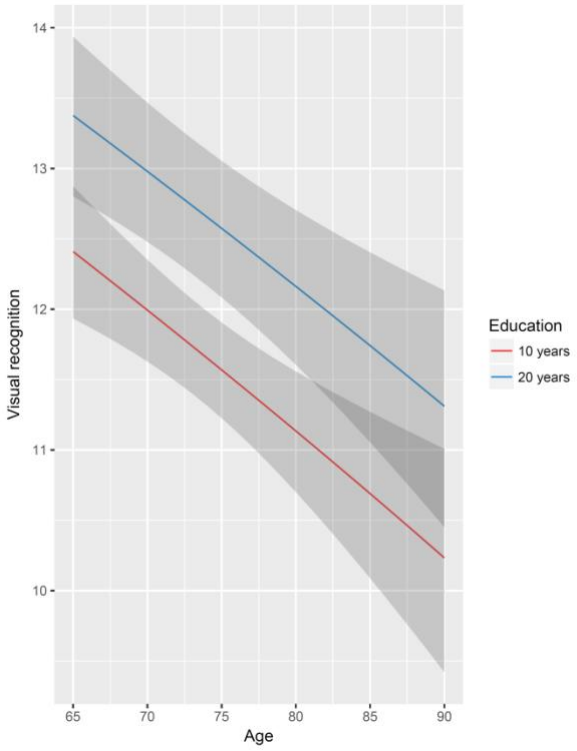


Picture learning and recognition:
Recalling images

	Description of question or task	Outcome measure	Units	Cognitive domain
Demographic and medical data				
Vision	Use of vision aids	Yes or no	%	Self-reporting
Hearing	Presence of hearing impairment or use of hearing aids	Yes or no	%	Self-reporting
Medication	Entry of the number of daily drugs	Either 0, 1–3, 4–7, or >7	%	Self-reporting
Age	Entry of the age in years	10–99	% of correct	Orientation
Education	Entry of the education in years	[illegible]	% of correct	Self-reporting
Language	German indicated as native language	1 out of 10 options	Not applicable	Self-reporting
	If the native language is not German: rating of the German language skills	None, basic, fluent	Not applicable	Self-reporting
Cognitive self-assessment				
	Answering 5 questions about the ability to cope with daily situations compared to 2 years ago: (1) memorizing new things; (2) remembering names; (3) multiple simultaneous tasks; (4) financial issues; (5) remembering appointments.	Likert scale from [illegible]	Points	Self-reporting
Temporal orientation				
Weekday	Entry of the weekday (Monday to Sunday)	Accuracy of performance	Correct or false	Orientation
Date	Entry of the day, month, and year	Accuracy of performance	Correct or false	Orientation
Subtests of cognitive functions				
Visual recognition	15-item multiple-choice version of the Benton Visual Retention Test. The examinee views each design for 10 seconds, and then chooses the correct design from a multiple choice of 4 displays.	Accuracy of performance	No of correct items	Visual memory
Picture learning and recognition	Recognition of 10 hand-drawings that were shown for 3 seconds each among 30 pictures at the end of the test. For each item, 2 pictures of the same semantic category serve as distractors.	Accuracy of performance	No of correct items	Episodic memory
Digit span	Repetition of an extending series of randomly presented digits using a number pad	Accuracy of performance	No of correct series	Attention span
Spatial span	Repetition of an extending series of blinking blocks	Accuracy of performance	No of correct series	Visuospatial attention
Reaction time and attention*	Tapping on the screen as fast as possible if a star appears (15 stars out of 15 other geometric forms)	Accuracy and speed of performance	No of correct hits, reaction time	Reaction time and attention
Trail Making Test A	Electronic version of the Trail Making Test A, limited in time (45 sec)	Accuracy and speed of performance	No of line connections/min	Psychomotor speed
Trail Making Test B	Electronic version of the Trail Making Test B, limited in time (60 sec)	Accuracy and speed of performance	No of line connections/min	Flexibility

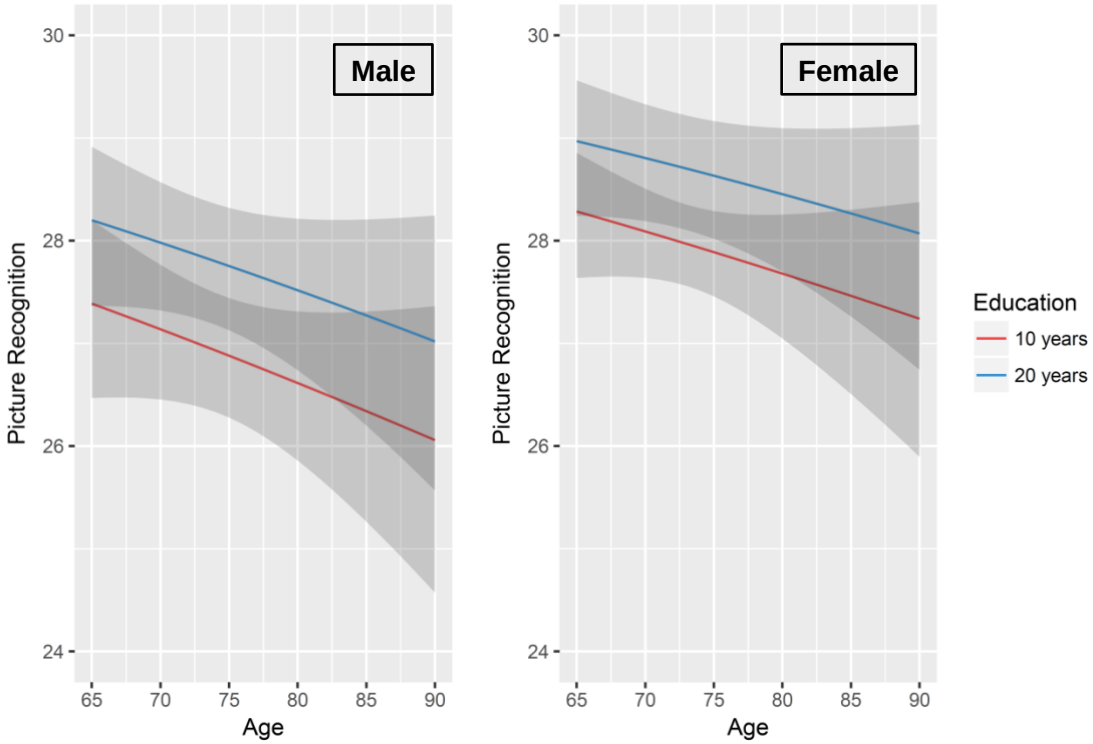
Supplemental digital content 3. Influence of demographic characteristics on CogCheck subtests.

Visual recognition



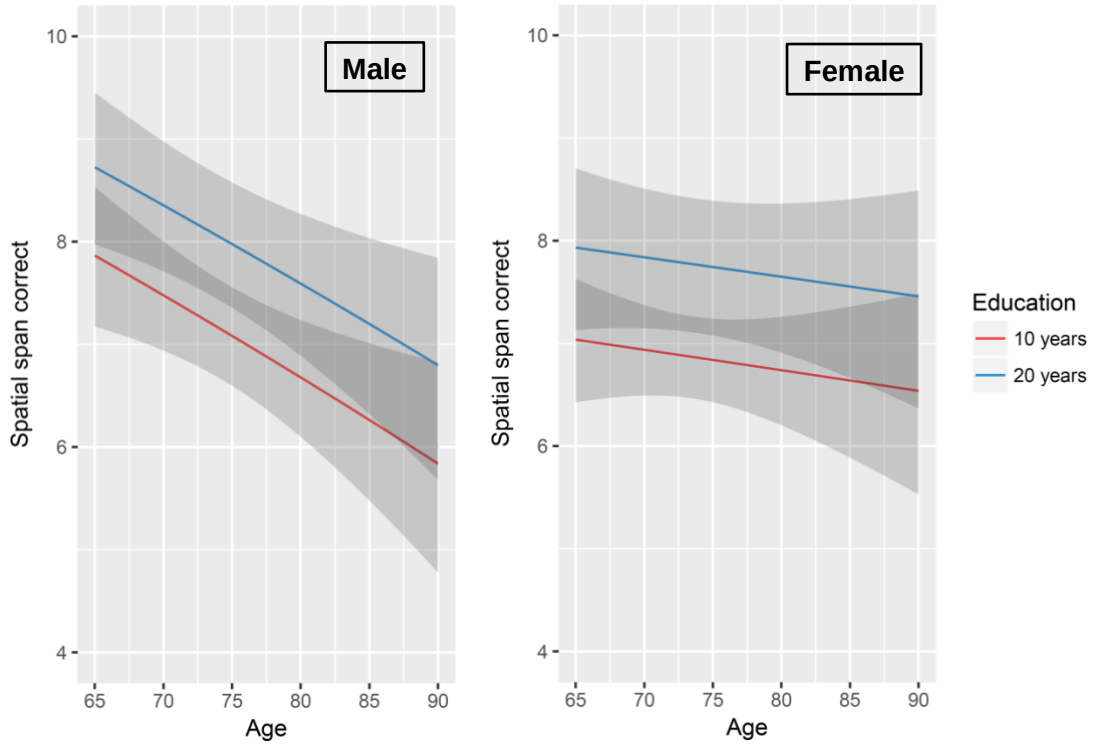
The best predictive model for the visual recognition score included age (β = -0.398, SE = 0.095, *P* < 0.001) and education (β = 0.479, SE = 0.171, *P* = 0.006), which explained a significant amount of variance of the visual recognition score, F = 12.71, *P* < 0.001, R^2 = 0.083 (adjusted R^2 = 0.077). In this model, older age, and lower education were associated with a lower visual recognition score.

Picture recognition



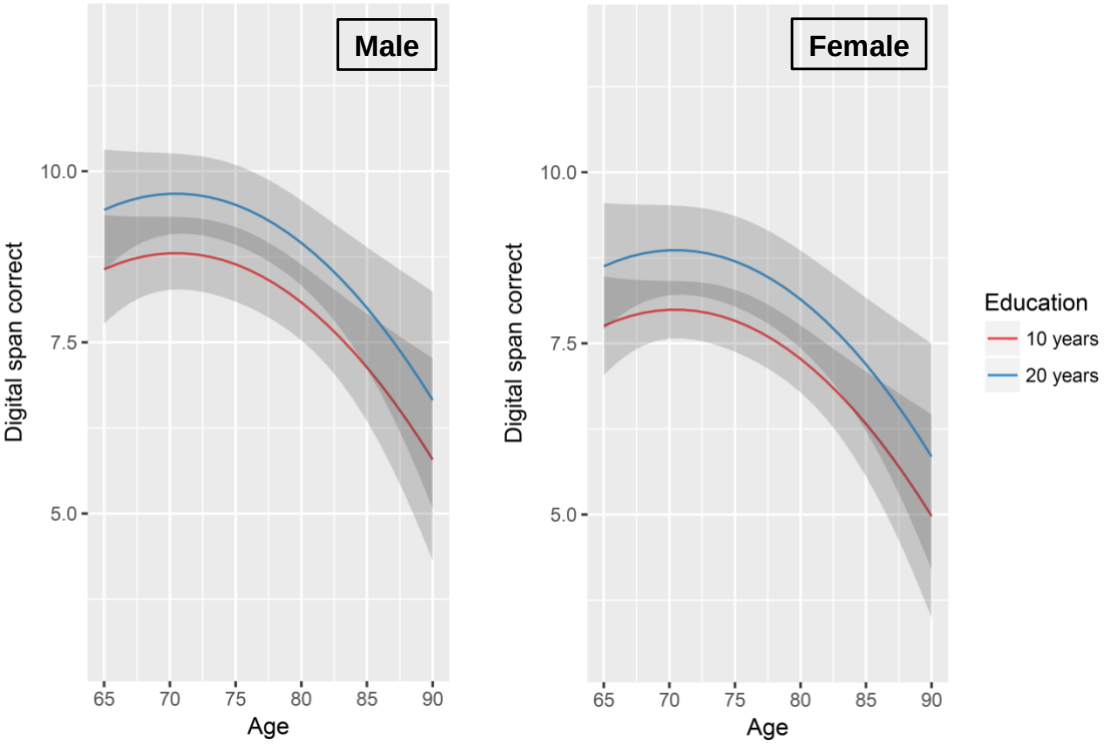
The best predictive model for the picture recognition score included age (β = -0.004, SE = 0.001, *P* = 0.004), education (β = 0.005, SE = 0.002, *P* = 0.026), gender (β = 0.056, SE = 0.014, *P* < 0.001), and a quadratic formula for age (β = 0.001, SE = 0.0002, *P* = 0.003), which explained a significant amount of variance of the picture recognition score, F = 7.56, *P* < 0.001, R^2 = 0.098 (adjusted R^2 = 0.085). In this model, lower education, and male gender were associated with a lower picture recognition score.

Spatial span



The best predictive model for the spatial span score included age (β = -0.245, SE = 0.083, *P* = 0.004), education (β = 0.141, SE = 0.121, *P* = 0.25), gender (β = -0.964, SE = 0.628, *P* = 0.13), the interaction of age and gender (β = 0.185, SE = 0.116, *P* = 0.11), and a quadratic formula for education (β = 0.050, SE = 0.030, *P* = 0.10), which explained a significant amount of variance of the spatial span score, F = 4.02, *P* = 0.002, R^2 = 0.068 (adjusted R^2 = 0.051). In this model, older age, and female gender were associated with a lower spatial span score.

Digit span



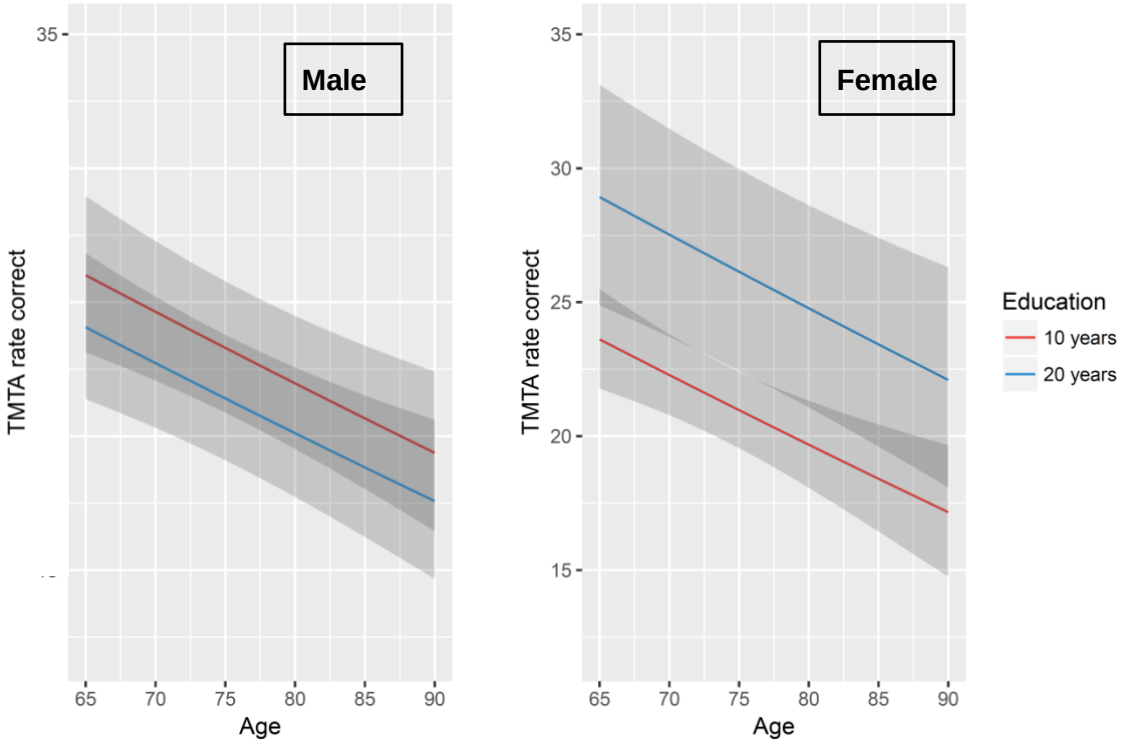
The best predictive model for the digit span score included age (β = -0.053, SE = 0.024, *P* = 0.030), education (β = 0.087, SE = 0.041, *P* = 0.036), gender (β = -0.811, SE = 0.239, *P* < 0.001), and a quadratic formula for age (β = -0.008, SE = 0.003, *P* = 0.020), which explained a significant amount of variance of the digit span score, F = 9.13, *P* < 0.001, R^2 = 0.117 (adjusted R^2 = 0.104). In this model, lower education, and female gender were associated with a lower digit span score.

Reaction time and attention

The initial assessment test included a reaction time and attention task. Participants were asked to tap on the screen as fast as possible if a star (n = 15) appeared in random order among other geometric forms (n = 15).

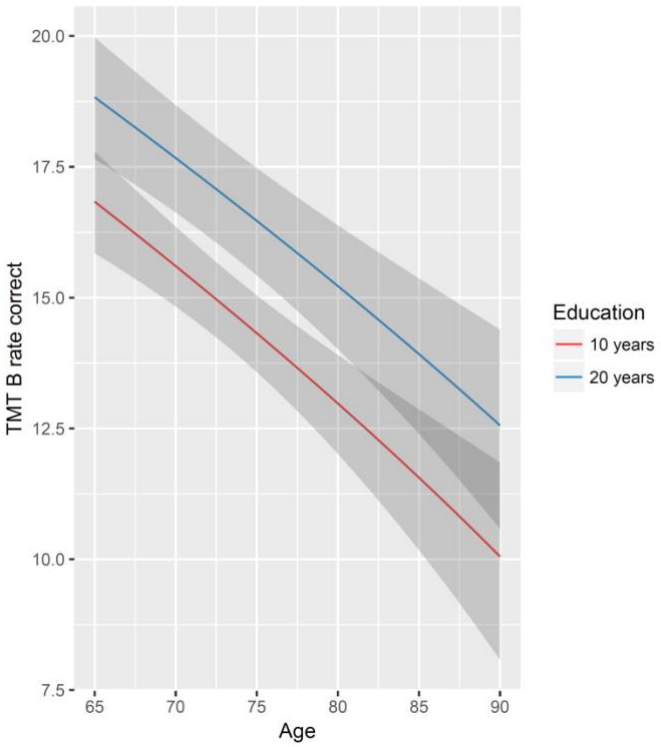
We found a ceiling effect for this task, and it was declared to be too easy (80% reached the maximum score of hits and 83% made no mistake). Moreover, the reaction time measurements were hampered by a programming error. As this precluded any standardization procedure, the reaction time and attention task was discarded from the assessment application.

Trail Making Test A



In Trail Making Test A (TMT-A), the ratio of correct line connections per minute was calculated to measure the accuracy of performance. The best predictive model for the TMT-A score included age (β = -0.091, SE = 0.023, $P < 0.001$), education (β = -0.136, SE = 0.079, $P = 0.086$), gender (β = 0.064, SE = 0.250, $P = 0.80$), the interaction of education and gender (β = 0.241, SE = 0.098, $P = 0.015$), and a quadratic formula for education (β = 0.026, SE = 0.014, $P = 0.068$), which explained a significant amount of variance of the TMT-A score, F = 4.3, $P < 0.001$, $R^2 = 0.073$ (adjusted $R^2 = 0.056$). In this model, older age was associated with a lower TMT-A score.

Trail Making Test B



In Trail Making Test B (TMT-B), the ratio of correct line connections per minute was calculated to measure the accuracy of performance. The best predictive model for the TMT-B score included age (β = -1.488, SE = 0.249, $P < 0.001$) and education (β = 1.264, SE = 0.451, $P = 0.005$), which explained a significant amount of variance of the TMT-B score, F = 21.9, $P < 0.001$, $R^2 = 0.138$ (adjusted $R^2 = 0.131$). In this model, older age and lower education were associated with a lower TMT-B score.

6. General discussion

In the presented studies, we have investigated the performance of cognitively healthy individuals in the German-version of a well-known screening tool for MCI (MoCA, study I) and in a newly developed self-administered computerized cognitive assessment (CogCheck, study III). Further, we have provided insight on the cognitive performance of patients with mild or major NCD in the MoCA, and we have introduced a new way to evaluate cognitive performance using two separate cut-offs (study II). Finally, we have discussed important heterogeneities between international normative studies which may inform future endeavors to create methodological guidelines in neuropsychology.

6.1 Diagnostic accuracy and two separate cut-offs

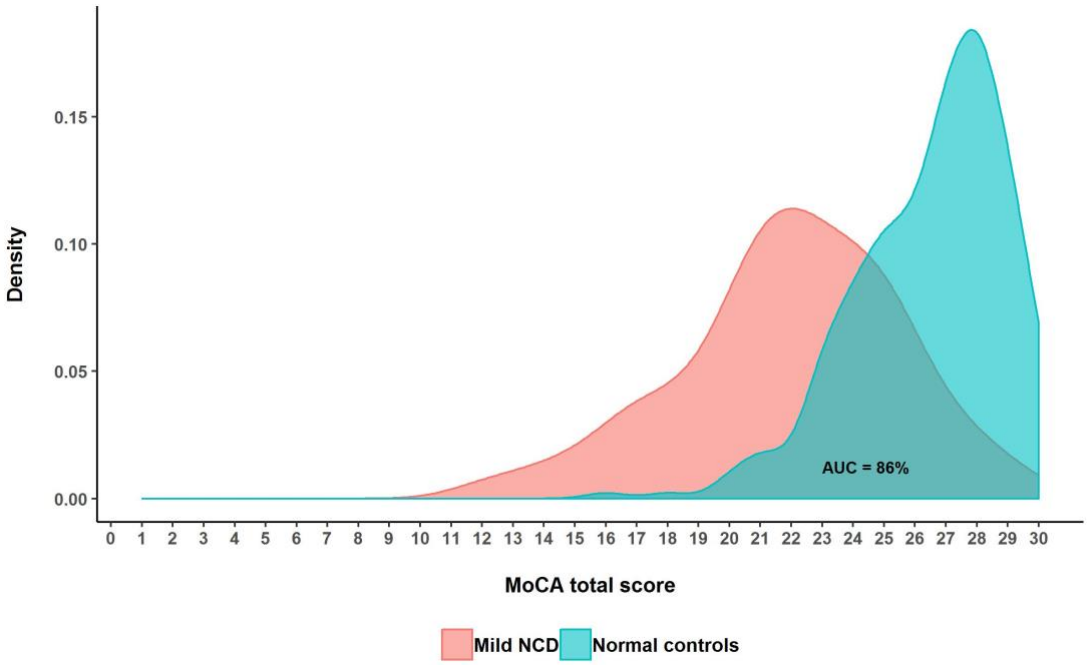
We have found demographic effects on cognitive performance in cognitively healthy individuals in the MoCA (study I), as well as in the CogCheck (study III). Consequently, we provided formulas to convert raw scores to demographically adjusted z-scores. Cognitively healthy individuals who are older and/or have lower educational attainment typically have lower cognitive performance and therefore are prone to false-positive test results. In study II, we tested this hypothesis and further assessed the diagnostic accuracy of the MoCA in consecutively referred Memory Clinic patients to distinguish patients with mild or major NCD from cognitive normal findings. As hypothesized, there were less false-positive test results when the MoCA z-score was applied to correct for demographic effects. However, using the MoCA z-score lead to decreased sensitivity to detect patients with mild or major NCD, which may be explained with age and education being important risk factors for dementia (Strauss et al., 2006). A follow-up study on the diagnostic accuracy of the CogCheck is currently in planning.

While the diagnostic accuracy of the MoCA was very high to distinguish patients with mild or major NCD from cognitive normal findings, the classification of mild NCD only versus cognitively healthy individuals was more challenging. Indeed, there was always an unsatisfying trade-off between sensitivity and specificity. In the context of neurodegeneration and cognitive impairment, both—false negative

and false-positive results—have undesirable implications for the patient as well as for the healthcare system, and there is no valid reason to favor one over the other.

Therefore, we have proposed a new way two evaluate cognitive performance on the MoCA by introducing two separate cut-offs in combination with a gray area. Specificity increases with lower scores, since healthy individuals typically have higher scores compared to patients. Inversely, higher scores are related to greater sensitivity, since patients typically have low scores. This is illustrated for the MoCA in Fig. 1. Instead of using one cut-off, which is always a trade-off between sensitivity and sensitivity, two separate cut-offs may be applied. A lower cut-off with high specificity and a higher cut-off with high sensitivity. Test results between these cut-off scores are represented by a gray area and require further examination and/or follow-up testing.

Fig. 1. Score distribution in patients with mild NCD and normal controls.



In Fig. 1, the score distribution of patients with mild NCD vs. normal controls is illustrated. The overlap between the score distributions corresponds to the discriminative accuracy (expressed in the area under the curve [AUC]). A test with smaller overlap has higher discriminative power than a test with greater overlap.

Lower MoCA scores are typically associated with more cognitive impairment. This information is lost, when a continuous variable like the MoCA is binarized. Considering the difficulties, the field is facing in identifying subtle cognitive decline, it is important to identify potential sources of information. Applying two separate cut-offs may be one way to make use of additional information provided by the MoCA. Of course, these findings are preliminary. Follow-up studies are required to test this method in other settings and on tools other than the MoCA. Another way to profit from continuous information of the MoCA may be the analysis on a sub-item level. Big data approaches and the use of machine learning algorithms have received increasing attention in clinical diagnostics (Sajda, 2006; Weakley, Williams, Schmitter-Edgecombe, & Cook, 2015). It is imaginable, that MoCA error-patterns are specific for patients with different pathologies and that these patterns might be decoded using multivariate approaches. Additionally, feature selection may reveal which sub-items are especially informative for the classification of healthy individuals versus patients. If less informative items could be removed from the MoCA, its administration time could be shortened. Computerized cognitive assessments may facilitate future data collection and data-driven insights to this aim.

6.2 Computerized cognitive assessment

In study III, we have investigated a novel computerized cognitive assessment tool (CogCheck) in cognitively healthy individuals. Compared to paper-pencil tests, computerized test batteries have many advantages. In neuropsychology, one strategy is the evaluation of qualitative aspects of cognitive assessment. This strategy, known as the Boston Process Approach, is interested in how a patient gets to a result, in addition to evaluating the correctness of the answer (Casaletto & Heaton, 2017; Libon, Swenson, Ashendorf, Bauer, & Bowers, 2013). With computerized assessments, such qualitative measures can easily be captured as a complement to a total score. For instance, reaction times, number and types of errors or self-corrections, and ways of proceeding can be assessed and evaluated. Moreover, the recognition of somatic symptoms, such as tremor, may also be included in a computerized tool to supplement cognitive measures. Digital measures of tremor relate very well to common clinical rating scales and even detected subtle abnormalities that are not captured in traditional clinical

assessments (Lipsmeier et al., 2018). Likewise, qualitative measures of cognition may provide valuable information in individuals who score within the normal range in standard cognitive assessments. Computerized assessments further yield the advantage of eliminating examiner-related bias, since the test items are always presented the exact same way. Moreover, assessments may be automatically adjusted in difficulty based on individual performance (Bauer et al., 2012). A neuropsychological assessment in the context of dementia diagnostics is usually performed just once or repeated only a few months later. Therefore, the gained insights are of cross-sectional nature and they may be biased if, for instance, a patient was very nervous at the time of testing. Moreover, the decline from a previous functional level may not be captured if an individual has a very high level of functioning. Computerized assessments may be administred repeatedly over a longer time-frame. This way, bias may be reduced by averaging the results and a possible decline becomes apparent.

Computerized cognitive assessments may additionally increase the efficiency of a cognitive evaluation. Indeed, results are available immediately after examination, and automatic scoring is less prone to errors. Fewer materials and less trained personnel are required, thereby reducing costs. It is imaginable, that patients may complete computerized assessments at home, which enhances the accessibility for people who have difficulties to travel to a clinician (Casaletto & Heaton, 2017). In current neuropsychological assessments, a standardized set of cognitive tests is typically combined with a hypothesis-driven approach, where additional tests are administered based on a specific referral question or based on a suspicion of impairment in a particular cognitive domain. Following this approach, a computerized assessment could serve as a first step evaluation in all patients, to gain a global impression of cognitive performance. Individually selected, specific tests may then be administered only in those patients, where more information is required. Such a multi-step approach on the level of specialized cognitive assessment may enhance efficiency.

While computerized cognitive assessments yield many advantages, there are challenges that need to be considered as well. It must be assured, that patients perform the test alone and without any additional help. Moreover, patients will typically be faced with their deficits during a cognitive assessment, which may affect the motivation to continue the evaluation. Neuropsychologists are trained to help motivate a patient in these situations. It should be evaluated, whether and to

what extent this lack of a motivating healthcare professional affects the use of self-administered computerized cognitive assessments. The absence of a trained professional may also have a negative impact on diagnostic conclusions, when individual characteristics that may impact cognitive performance (e.g., lack of motivation, sensorimotor impairment, very low educational attainment) are not considered for test interpretation (American Academy of Clinical Neuropsychology, 2007; Bauer et al., 2012). Further, important ethical and data safety considerations should be considered, especially if a computerized tool is connected to the Internet. Finally, the psychometric properties of any cognitive assessment tool need to be evaluated (Casaletto & Heaton, 2017) and normative values are required. The reliability and validity of CogCheck will therefore have to be investigated in follow-up studies before its application in a clinical setting. Further, other important questions need to be addressed for the successful implementation of computerized cognitive assessment tools in clinical neuropsychology. To what extent are the psychometric properties altered in a computerized tool? How do technical changes, on a current test version, alter the psychometric properties of the updated version? Are normative values, that have been developed in one setting, transferable to other settings? How well do normative values from one culture translate to another?

6.3 Definition of cognitive health and selection of patient groups

Whether normative values that have been developed in one country or in a specific setting are representable for other countries/settings is a central question, that the field of neuropsychology needs to address (Casaletto & Heaton, 2017). Indeed, it is expensive to conduct normative studies and it is not possible to create new normative values in every language and every setting. Additionally, norms that have been created in older generations may not translate to newer generations and may need to be renewed. As outlined in the previous chapter, computerized cognitive assessments are becoming more popular. However, technical developments underlie rapid changes and new technical advances are usually quickly outdated and replaced by newer technologies. This rapidly changing environment is a main challenge that needs to be addressed in neuropsychology. It is simply impossible to keep up with the technical advances if normative values should be gathered for every new version of a tool.

Therefore, the field would benefit from the knowledge whether and to what extent norms are generalizable across different cultures and settings. However, answering this question requires a clear consensus and guidelines on (a) the requirements of a normative sample and (b) the methodology that is used to create norms. Only then, the potential differences in cognitive performance between cultures/settings may truly be attributable to these factors. Otherwise, an alternative explanation for the observed differences may be the inhomogeneity in subject selection and methodology.

In a side project (Thomann, Goettel, Hessler, et al., 2018), we have compared MoCA scores from culturally similar study centers (i.e. the Basel normative sample [Thomann, Goettel, Monsch, et al., 2018] vs. a normative sample from Munich, Germany/Vienna, Austria) and found significantly different mean MoCA scores ($p < .001$) between the two study centers. In a regression-model, lower MoCA scores were associated with higher age ($p = .005$), lower education ($p = .003$), male sex ($p = .018$), and study center (Munich/Vienna; $p = .003$). There is no reason to assume, that cognitive performance in Swiss participants would differ from the performance in the culturally- and language-related German and Austrian participants. Moreover, there were no interactions between study center and demographic characteristics that could explain the difference in mean MoCA scores. Therefore, we compared the samples regarding their inclusion/exclusion criteria and found important differences, on how cognitive health was defined (CERAD total score and MMSE cut-off in Basel vs. subjective memory complaints in Munich/Vienna). These dissimilarities in exclusion criteria may explain the observed differences in mean MoCA total scores.

Likewise, we have found differences in mean MoCA total scores between international MoCA normative studies (study I). Again, there were important dissimilarities with respect to the methodology and the sample characteristics between these studies. Notably, the definition of cognitive health was largely inconsistent. There seems to be no clear consensus regarding the selection of cognitively healthy individuals for a normative study, and opinions diverge in whether a population-based approach or a sample with indisputable cognitively healthy individuals to create robust norms should be favored (Casaletto & Heaton, 2017; Martin et al., 2017; Strauss et al., 2006). Some argue, that cognitively healthy individuals constitute a sample of “superhealthy” individuals who are not representative for the general population. In study II, we therefore compared our

healthy normative sample to cognitive normal findings from clinical routine and we have found no difference in the MoCA performance between these two groups. This suggests that a purposely-recruited cognitively healthy sample is representable for cognitive normal findings from clinical routine. However, this result should be completed by longitudinal data in future investigations.

The field of AD research in general could profit from (a) a clear definition of cognitive health and (b) revised criteria for cognitive impairment. When actuarial neuropsychological criteria are used to define MCI, more patients and healthy individuals are correctly classified (Bondi et al., 2014; Edmonds et al., 2018; Edmonds et al., 2016). This is essential for clinical care, to ensure that only patients receive further examinations and treatments. Moreover, such criteria are important in clinical research, since including the correct participants in a study (i.e., truly patients/truly healthy individuals) is fundamental. Indeed, the erroneous inclusion of patients as healthy participants in studies on diagnostic tests lowers the mean performance of the group, which in term leads to lower cut-off scores and may ultimately decrease sensitivity. Likewise, including healthy individuals as patients may severely bias study results and even mask treatment effects (Edmonds et al., 2018). Moreover, wrong conclusions about the pathogenesis and the prediction of the disease course may be drawn.

7. Outlook

The changes that arise from the demographical development, combined with the digitalization, yields challenges, as well as opportunities for our society and the healthcare system. Neuropsychology has important contributions to offer in facing these challenges and by modernizing the way we assess cognition.

The rapidly growing segment of older people requires fast and accurate diagnostic methods to identify patients with cognitive disorders. The field of neuropsychology needs to incorporate new insights of brain-behavior relationships into the current clinical practice to address this. Further, new opportunities that arise form technical advances should be seized to keep the neuropsychological instruments up to date. Computerized cognitive assessments yield the advantage of highly time- and resource-efficient examinations. They may additionally facilitate the development of normative values, since computerized assessments may be made available as a mobile application in large-scale studies to a broad range of individuals without the presence of a trained examiner (e.g., GameChanger study [Alzheimer's Society, 2019]). Such mobile applications have been used in other disease areas for daily symptom assessment with active and passive measures (Lipsmeier et al., 2018). Therefore, the digitalization offers huge opportunities to gather information from many different sources that may open new insights on healthy or pathological processes and which might be translated into so-called digital biomarkers. Hypothesis-free cluster analyses may then be applied to investigate, whether there are specific groups of patients with individual characteristics, which need to be considered to enhance diagnostic accuracy.

The approach of precision medicine increasingly gains popularity in the context of treatment and prevention of diseases. The consideration of individual differences may also be meaningful for the diagnostic process. In future studies, it may therefore be interesting to investigate whether other characteristics than education, age, or sex should be considered when cognitive performance is evaluated. For instance, cognitive performance is often discussed in the context of cognitive reserve (Stern, 2009), which postulates a protection against the harmful effects of neurodegeneration. Cognitive reserve may be modulated over the life course through cognitive activity or a complex occupation. Current research aims at finding a reliable marker for cognitive reserve, which may be incorporated in cognitive assessments as a co-variate. Moreover, some populations with above-normal

behavioral phenotypes show altered brain connections (Brauchli, Leipold, & Jancke, 2019) and enhanced cognitive abilities (Mealor, Simner, & Ward, 2019) compared to control groups, which in term could lead to performance differences in cognitive assessments. In conclusion, the field of neuropsychology has some interesting questions left to answer, which may help increase our understanding of brain-behavior-relationships and of possible alterations of these interactions by normal ageing. Such knowledge is central to further gain new insights into pathological ageing mechanisms. Ultimately, we need to find ways to disentangle the overlapping distributions between patients and healthy individuals. This may be achieved by the incorporation of new information, the use of new technologies, and the consideration of individual differences. One promising approach might be the combination of multimodal information with decision tree algorithms. There may be ways to screen for subtle symptoms with very few effort or even with passive monitoring through digital wearables or smartphones. Complemented with known characteristics of an individual, learning algorithms may automatically analyze which further diagnostic steps are indicated. However, the approach of precision medicine may also require clear methodological guidelines and internationally harmonized definitions for cognitive health and cognitive impairment. Otherwise, variability that arises from methodological dissimilarities may be misinterpreted as individual differences, or individual differences may be masked.

In conclusion, efficiency and diagnostic accuracy can be enhanced by applying automated, self-administrated, and simple diagnostic tools for the general population in combination with more specific further examinations that are tailored to the needs and characteristics of a specific individual.

8. References

Abou-Mrad, F., Chelune, G., Zamrini, E., Tarabey, L., Hayek, M., & Fadel, P. (2017). Screening for dementia in Arabic: normative data from an elderly Lebanese sample. *Clinical Neuropsychologist*, 1-19. doi:10.1080/13854046.2017.1288270

Aldecoa, C., Bettelli, G., Bilotta, F., Sanders, R. D., Audisio, R., Borozdina, A., . . . MacLullich, A. (2017). European Society of Anaesthesiology evidence-based and consensus-based guideline on postoperative delirium. *Eur J Anaesthesiol, 34*(4), 192-214.

Alzheimer's Society. GameChanger. Retrieved from https://www.alzheimers.org.uk/research/play-your-part/gamechanger

American Academy of Clinical Neuropsychology (AACN) practice guidelines for neuropsychological assessment and consultation. (2007). *Clin Neuropsychol, 21*(2), 209-231. doi:10.1080/13825580601025932

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. (2013). Arlington, VA: American Psychiatric Association.

Anyiam, F. (2018). *Risk Assessment for Postoperative Delirium: Implementation of a Self-administered Tablet Computer-based Clinical Screening Tool in a Clinical Setting.* (MD), University of Basel, Basel, Switzerland.

Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Clin Neuropsychol, 26*(2), 177-196.

Benton, A. L. (1972). Abbreviated versions of the Visual Retention Test. *J Psychol, 80*(2d Half), 189-192. doi:10.1080/00223980.1972.9924794

Berres, M., Zehnder, A., Blasi, S., & Monsch, A. U. (2008). Evaluation of diagnostic scores with adjustment for covariates. *Stat Med, 27*(10), 1777-1790. doi:10.1002/sim.3120

Bondi, M. W., Edmonds, E. C., Jak, A. J., Clark, L. R., Delano-Wood, L., McDonald, C. R., . . . Salmon, D. P. (2014). Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. *J Alzheimers Dis, 42*(1), 275-289. doi:10.3233/JAD-140276

Borland, E., Nagga, K., Nilsson, P. M., Minthon, L., Nilsson, E. D., & Palmqvist, S. (2017). The Montreal Cognitive Assessment: Normative Data from a Large Swedish Population-Based Cohort. *Journal of Alzheimer's Disease, 59*(3), 893-901. doi:10.3233/JAD-170203

Brauchli, C., Leipold, S., & Jancke, L. (2019). Univariate and multivariate analyses of functional networks in absolute pitch. *Neuroimage, 189*, 241-247. doi:10.1016/j.neuroimage.2019.01.021

Brown, C., & Deiner, S. (2016). Perioperative cognitive protection. *Br J Anaesth, 117*(suppl 3), iii52-iii61. doi:10.1093/bja/aew361

Brown, M. D., & Reeves, M. J. (2003). Evidence-based emergency medicine/skills for evidence-based emergency care. Interval likelihood ratios: another advantage for the evidence-based diagnostician. *Ann Emerg Med, 42*(2), 292-297.

Brunetti, R., Del Gatto, C., & Delogu, F. (2014). eCorsi: implementation and testing of the Corsi block-tapping task for digital tablets. *Front Psychol, 5*, 939.

Burckhardt, A. C. (2014). *Das Delir Risiko Profil: Ein computerisiertes Untersuchungsinstrument zur präoperativen Erfassung des postoperativen Delir-Risikos.* (MD), University of Basel, Basel, Switzerland.

Carson, N., Leach, L., & Murphy, K. J. (2018). A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. *Int J Geriatr Psychiatry, 33*(2), 379-388. doi:10.1002/gps.4756

Casaletto, K. B., & Heaton, R. K. (2017). Neuropsychological Assessment: Past and Future. *J Int Neuropsychol Soc, 23*(9-10), 778-790. doi:10.1017/S1355617717001060

Chapman, K. R., Bing-Canar, H., Alosco, M. L., Steinberg, E. G., Martin, B., Chaisson, C., . . . Stern, R. A. (2016). Mini Mental State Examination and Logical Memory scores for entry into Alzheimer's disease trials. *Alzheimers Res Ther, 8*, 9. doi:10.1186/s13195-016-0176-z

Ciesielska, N., Sokolowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A., & Kedziora-Kornatowska, K. (2016). Is the Montreal Cognitive Assessment (MoCA) test better suited than the Mini-Mental State Examination (MMSE) in mild cognitive impairment (MCI) detection among people aged over 60? Meta-analysis. *Psychiatr Pol, 50*(5), 1039-1052. doi:10.12740/PP/45368

Clarfield, A. M. (2003). The decreasing prevalence of reversible dementias: an updated meta-analysis. *Archives of Internal Medicine, 163*(18), 2219-2229. doi:10.1001/archinte.163.18.2219

CogState Ltd. COGNIGRAM. Retrieved from https://www.cognigram.com/

Conti, S., Bonazzi, S., Laiacona, M., Masina, M., & Coralli, M. V. (2015). Montreal Cognitive Assessment (MoCA)-Italian version: regression based norms and equivalent scores. *Neurological Sciences, 36*(2), 209-214. doi:10.1007/s10072-014-1921-3

Crook, T. H., Kay, G. G., & Larrabee, G. L. (2009). Computer-Based Cognitive Testing. In I. Grant & K. M. Adams (Eds.), *Neuropsychological Assessment of Neuropsychiatric and Neuromedical Disorders* (3rd ed.). Oxford, UK: Oxford University Press.

Dasgupta, M., & Dumbrell, A. C. (2006). Preoperative risk assessment for delirium after noncardiac surgery: a systematic review. *J Am Geriatr Soc, 54*(10), 1578-1589.

Davis, D. H., Creavin, S. T., Yip, J. L., Noel-Storr, A. H., Brayne, C., & Cullum, S. (2015). Montreal Cognitive Assessment for the diagnosis of Alzheimer's disease and other dementias. *Cochrane Database Syst Rev*(10), CD010775. doi:10.1002/14651858.CD010775.pub2

Edmonds, E. C., Ard, M. C., Edland, S. D., Galasko, D. R., Salmon, D. P., & Bondi, M. W. (2018). Unmasking the benefits of donepezil via psychometrically precise identification of mild cognitive impairment: A secondary analysis of the ADCS vitamin E and donepezil in MCI study. *Alzheimers Dement (N Y), 4*, 11-18. doi:10.1016/j.trci.2017.11.001

Edmonds, E. C., Delano-Wood, L., Jak, A. J., Galasko, D. R., Salmon, D. P., Bondi, M. W., & Alzheimer's Disease Neuroimaging, I. (2016). "Missed" Mild Cognitive Impairment: High False-Negative Error Rate Based on Conventional Diagnostic Criteria. *J Alzheimers Dis, 52*(2), 685-691. doi:10.3233/JAD-150986

Ehrensperger, M. M., Berres, M., Taylor, K. I., & Monsch, A. U. (2010). Early detection of Alzheimer's disease with a total score of the German CERAD. *J Int Neuropsychol Soc, 16*(5), 910-920. doi:10.1017/S1355617710000822

Ehrensperger, M. M., Taylor, K. I., Berres, M., Foldi, N. S., Dellenbach, M., Bopp, I., . . . Monsch, A. U. (2014). BrainCheck - a very brief tool to detect incipient cognitive decline: optimized case-finding combining patient- and informant-based data. *Alzheimers Res Ther, 6*(9), 69. doi:10.1186/s13195-014-0069-y

Erlanger, D. M., Kaushik, T., Broshek, D., Freeman, J., Feldman, D., & Festa, J. (2002). Development and validation of a web-based screening tool for monitoring cognitive status. *J Head Trauma Rehabil, 17*(5), 458-476.

Evered, L. A. (2017). Predicting delirium: are we there yet? *Br J Anaesth, 119*(2), 281-283. doi:10.1093/bja/aex082

Fazeli, P. L., Ross, L. A., Vance, D. E., & Ball, K. (2013). The relationship between computer experience and computerized cognitive test performance among older adults. *J Gerontol Ser B-Psychol Sci Soc Sci, 68*(3), 337-346.

Florkowski, C. M. (2008). Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: communicating the performance of diagnostic tests. *Clin Biochem Rev, 29 Suppl 1*, S83-87.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*(3), 189-198.

Freitas, S., Simoes, M. R., Alves, L., & Santana, I. (2011). Montreal Cognitive Assessment (MoCA): normative study for the Portuguese population. *Journal of Clinical and Experimental Neuropsychology, 33*(9), 989-996. doi:10.1080/13803395.2011.589374

Freter, S. H., Dunbar, M. J., MacLeod, H., Morrison, M., MacKnight, C., & Rockwood, K. (2005). Predicting post-operative delirium in elective orthopaedic patients: the Delirium Elderly At-Risk (DEAR) instrument. *Age Ageing, 34*(2), 169-171. doi:10.1093/ageing/afh245

Frisoni, G. B., Boccardi, M., Barkhof, F., Blennow, K., Cappa, S., Chiotis, K., . . . Winblad, B. (2017). Strategic roadmap for an early diagnosis of Alzheimer's disease based on biomarkers. *Lancet Neurol, 16*(8), 661-676. doi:10.1016/S1474-4422(17)30159-X

Fu, C., Jin, X., Chen, B., Xue, F., Niu, H., Guo, R., . . . Zhang, Y. (2017). Comparison of the Mini-Mental State Examination and Montreal Cognitive Assessment executive subtests in detecting post-stroke cognitive impairment. *Geriatr Gerontol Int, 17*(12), 2329-2335. doi:10.1111/ggi.13069

Goldenberg, G., Kiselev, P., Bharathan, T., Baccash, E., Gill, L., Madhav, V., . . . Thakur, C. (2006). Predicting post-operative delirium in elderly patients undergoing surgery for hip fracture. *Psychogeriatrics, 6*(2), 43-48.

Habibzadeh, F., Habibzadeh, P., & Yadollahie, M. (2016). On determining the most appropriate test cut-off value: the case of tests with continuous results. *Biochem Med (Zagreb), 26*(3), 297-307. doi:10.11613/BM.2016.034

Hall, M. J., DeFrances, C. J., Williams, S. N., Golosinskiy, A., & Schwartzman, A. (2010). National Hospital Discharge Survey: 2007 summary. *Natl Health Stat Report*(29), 1-24.

Inouye, S. K., Bogardus, S. T., Jr., Charpentier, P. A., Leo-Summers, L., Acampora, D., Holford, T. R., & Cooney, L. M., Jr. (1999). A multicomponent intervention to prevent delirium in hospitalized older patients. *N Engl J Med, 340*(9), 669-676. doi:10.1056/NEJM199903043400901

Inouye, S. K., Viscoli, C. M., Horwitz, R. I., Hurst, L. D., & Tinetti, M. E. (1993). A predictive model for delirium in hospitalized elderly medical patients based on admission characteristics. *Ann Intern Med, 119*(6), 474-481.

Inouye, S. K., Westendorp, R. G., & Saczynski, J. S. (2014). Delirium in elderly people. *Lancet, 383*(9920), 911-922.

Jennen-Steinmetz, C., & Wellek, S. (2005). A new approach to sample size calculation for reference interval studies. *Stat Med, 24*(20), 3199-3212.

Joddrell, P., & Astell, A. J. (2016). Studies involving people with dementia and touchscreen technology: a literature review. *JMIR Rehabil Assist Technol, 3*(2), e10. doi:10.2196/rehab.5788.

Jones, R. N., Marcantonio, E. R., Saczynski, J. S., Tommet, D., Gross, A. L., Travison, T. G., . . . Cizginer, S. (2016). Preoperative cognitive performance dominates risk for delirium among older adults. *J Geriatr Psychiatry Neurol, 29*(6), 320-327.

Jones, R. N., Yang, F. M., Zhang, Y., Kiely, D. K., Marcantonio, E. R., & Inouye, S. K. (2006). Does educational attainment contribute to risk for delirium? A potential role for cognitive reserve. *J Gerontol Ser A-Biol Sci Med Sci, 61*(12), 1307-1311.

Kalisvaart, K. J., Vreeswijk, R., de Jonghe, J. F., van der Ploeg, T., van Gool, W. A., & Eikelenboom, P. (2006). Risk factors and prediction of postoperative delirium in elderly hip-surgery patients: implementation and validation of a medical risk factor model. *J Am Geriatr Soc, 54*(5), 817-822. doi:10.1111/j.1532-5415.2006.00704.x

Kenny, R. A., Coen, R. F., Frewen, J., Donoghue, O. A., Cronin, H., & Savva, G. M. (2013). Normative Values of Cognitive and Physical Function in Older Adults:

Findings from The Irish Longitudinal Study on Ageing. *Journal of the American Geriatrics Society, 61*(s2), S279-S290. doi:10.1111/jgs.12195

Konstantopoulos, K., Vogazianos, P., & Doskas, T. (2016). Normative Data of the Montreal Cognitive Assessment in the Greek Population and Parkinsonian Dementia. *Archives of Clinical Neuropsychology, 31*(3), 246-253. doi:10.1093/arclin/acw002

Kopecek, M., Stepankova, H., Lukavsky, J., Ripova, D., Nikolai, T., & Bezdicek, O. (2017). Montreal cognitive assessment (MoCA): Normative data for old and very old Czech adults. *Applied Neuropsychology: Adult, 24*(1), 23-29. doi:10.1080/23279095.2015.1065261

Lam, B., Middleton, L. E., Masellis, M., Stuss, D. T., Harry, R. D., Kiss, A., & Black, S. E. (2013). Criterion and convergent validity of the Montreal cognitive assessment with screening and standardized neuropsychological testing. *J Am Geriatr Soc, 61*(12), 2181-2185. doi:10.1111/jgs.12541

Larouche, E., Tremblay, M. P., Potvin, O., Laforest, S., Bergeron, D., Laforce, R., . . . Hudon, C. (2016). Normative Data for the Montreal Cognitive Assessment in Middle-Aged and Elderly Quebec-French People. *Archives of Clinical Neuropsychology*. doi:10.1093/arclin/acw076

Lee, A., Mu, J. L., Joynt, G. M., Chiu, C. H., Lai, V. K. W., Gin, T., & Underwood, M. J. (2017). Risk prediction models for delirium in the intensive care unit after cardiac surgery: a systematic review and independent external validation. *Br J Anaesth, 118*(3), 391-399. doi:10.1093/bja/aew476

Libon, D. J., Swenson, R., Ashendorf, L., Bauer, R. M., & Bowers, D. (2013). Edith Kaplan and the Boston Process Approach. *Clin Neuropsychol, 27*(8), 1223-1233. doi:10.1080/13854046.2013.833295

Lijmer, J. G., Mol, B. W., Heisterkamp, S., Bonsel, G. J., Prins, M. H., van der Meulen, J. H., & Bossuyt, P. M. (1999). Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA, 282*(11), 1061-1066.

Lipsmeier, F., Taylor, K. I., Kilchenmann, T., Wolf, D., Scotland, A., Schjodt-Eriksen, J., . . . Lindemann, M. (2018). Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. *Mov Disord, 33*(8), 1287-1297. doi:10.1002/mds.27376

Long, L. S., Shapiro, W. A., & Leung, J. M. (2012). A brief review of practical preoperative cognitive screening tools. *Can J Anaesth, 59*(8), 798-804. doi:10.1007/s12630-012-9737-1

López-Ratón, M., Rodríguez-Álvarez, M. X., Cadarso-Suárez, C., & Gude-Sampedro, F. (2014). OptimalCutpoints: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests. *2014, 61*(8), 36. doi:10.18637/jss.v061.i08

Lu, J., Li, D., Li, F., Zhou, A., Wang, F., Zuo, X., . . . Jia, J. (2011). Montreal Cognitive Assessment in Detecting Cognitive Impairment in Chinese Elderly Individuals: A Population-Based Study. *Journal of Geriatric Psychiatry and Neurology, 24*(4), 184-190. doi:10.1177/0891988711422528

Malek-Ahmadi, M., Powell, J. J., Belden, C. M., O'Connor, K., Evans, L., Coon, D. W., & Nieri, W. (2015). Age- and education-adjusted normative data for the Montreal Cognitive Assessment (MoCA) in older adults age 70-99. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition, 22*(6), 755-761. doi:10.1080/13825585.2015.1041449

Marcantonio, E. R., Goldman, L., Mangione, C. M., Ludwig, L. E., Muraca, B., Haslauer, C. M., . . . et al. (1994). A clinical prediction rule for delirium after elective noncardiac surgery. *JAMA, 271*(2), 134-139.

Martin, P. K., Schroeder, R. W., & Baade, L. E. (2017). A tale of two norms: the impact of normative sample selection criteria on standardized scores in older adults. *Clin Neuropsychol, 31*(6-7), 1204-1218. doi:10.1080/13854046.2017.1349182

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr., Kawas, C. H., . . . Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement, 7*(3), 263-269. doi:10.1016/j.jalz.2011.03.005

Mealor, A. D., Simner, J., & Ward, J. Does synaesthesia protect against age-related memory loss? *Journal of Neuropsychology, 0*(0). doi:10.1111/jnp.12182

Mohanty, S., Rosenthal, R. A., Russell, M. M., Neuman, M. D., Ko, C. Y., & Esnaola, N. F. (2016). Optimal perioperative management of the geriatric patient: a best practices guideline from the American College of Surgeons NSQIP and the American Geriatrics Society. *J Am Coll Surg, 222*(5), 930-947.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., . . . Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology, 39*(9), 1159-1165.

Nadelson, M. R., Sanders, R. D., & Avidan, M. S. (2014). Perioperative cognitive trajectory in adults. *Br J Anaesth, 112*(3), 440-451. doi:10.1093/bja/aet420

Narazaki, K., Nofuji, Y., Honda, T., Matsuo, E., Yonemoto, K., & Kumagai, S. (2013). Normative data for the montreal cognitive assessment in a Japanese community-dwelling older population. *Neuroepidemiology, 40*(1), 23-29. doi:10.1159/000339753

Nasreddine, Z. S., Phillips, N., & Chertkow, H. (2012). Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology, 78*(10), 765-766; author reply 766. doi:10.1212/01.wnl.0000413072.54070.a3

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699.

National Clinical Guideline Centre (2010). National Institute for Health and Clinical Excellence: Guidance *Delirium: Diagnosis, Prevention and Management*. London: Royal College of Physicians (UK)

NIH Toolbox. Retrieved from http://www.healthmeasures.net/explore-measurement-systems/nih-toolbox

Noel-Storr, A. H., McCleery, J. M., Richard, E., Ritchie, C. W., Flicker, L., Cullum, S. J., . . . McShane, R. (2014). Reporting standards for studies of diagnostic test accuracy in dementia: The STARDdem Initiative. *Neurology, 83*(4), 364-373. doi:10.1212/WNL.0000000000000621

O'Connell, M. E., & Tuokko, H. (2010). Age corrections and dementia classification accuracy. *Arch Clin Neuropsychol, 25*(2), 126-138. doi:10.1093/arclin/acp111

Ozer, S., Young, J., Champ, C., & Burke, M. (2016). A systematic review of the diagnostic test accuracy of brief cognitive tests to detect amnestic mild cognitive impairment. *Int J Geriatr Psychiatry, 31*(11), 1139-1150. doi:10.1002/gps.4444

Pereiro, A. X., Ramos-Lema, S., Lojo-Seoane, C., Guàrdia-Olmos, J., Facal-Mayo, D., & Juncos-Rabadán, O. (2017). Normative data for the Montreal Cognitive Assessment (MOCA) in a Spanish sample of community-dweller adults. *European Geriatric Medicine*.

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *J Intern Med, 256*(3), 183-194. doi:10.1111/j.1365-2796.2004.01388.x

Petersen, R. C., Lopez, O., Armstrong, M. J., Getchius, T. S. D., Ganguli, M., Gloss, D., . . . Rae-Grant, A. (2017). Practice guideline update summary: Mild cognitive impairment. *Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology*. doi:10.1212/wnl.0000000000004826

Pocklington, C., Gilbody, S., Manea, L., & McMillan, D. (2016). The diagnostic accuracy of brief versions of the Geriatric Depression Scale: a systematic review and meta-analysis. *Int J Geriatr Psychiatry, 31*(8), 837-857. doi:10.1002/gps.4407

Prince, M., Bryce, R., & Ferri, C. (2011). World Alzheimer Report 2011: The benefits of early diagnosis and intervention. Retrieved from https://www.alz.co.uk/research/world-report-2011

Prince, M., Comas-Herrera, A., Knapp, M., Guerchet, M., & Karagiannidou, M. (2016). World Alzheimer Report 2016. Improving healthcare for people living with dementia: Coverage, quality and costs now and in the future. Retrieved from https://www.alz.co.uk/research/world-report-2016

Roalf, D. R., Moberg, P. J., Xie, S. X., Wolk, D. A., Moelter, S. T., & Arnold, S. E. (2013). Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimers Dement, 9*(5), 529-537. doi:10.1016/j.jalz.2012.10.001

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dement Geriatr Cogn Disord, 5*(5), 266-281.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Muller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12*, 77. doi:10.1186/1471-2105-12-77

Rossetti, H. C., Lacritz, L. H., Cullum, C. M., & Weiner, M. F. (2011). Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology, 77*(13), 1272-1275. doi:10.1212/WNL.0b013e318230208a

Rossetti, H. C., Lacritz, L. H., Hynan, L. S., Cullum, C. M., Van Wright, A., & Weiner, M. F. (2017). Montreal Cognitive Assessment Performance among Community-Dwelling African Americans. *Arch Clin Neuropsychol, 32*(2), 238-244. doi:10.1093/arclin/acw095

Sajda, P. (2006). Machine learning for detection and diagnosis of disease. *Annu Rev Biomed Eng, 8*, 537-565. doi:10.1146/annurev.bioeng.8.061505.095802

Salthouse, T. A., & Fristoe, N. M. (1995). Process analysis of adult age effects on a computer-administered Trail Making Test. *Neuropsychology, 9*(4), 518.

Sanders, R. D., Pandharipande, P. P., Davidson, A. J., Ma, D., & Maze, M. (2011). Anticipating and managing postoperative delirium and cognitive decline in adults. *BMJ, 343*, d4331. doi:10.1136/bmj.d4331

Santangelo, G., Siciliano, M., Pedone, R., Vitale, C., Falco, F., Bisogno, R., . . . Trojano, L. (2015). Normative data for the Montreal Cognitive Assessment in an Italian population sample. *Neurological Sciences, 36*(4), 585-591. doi:10.1007/s10072-014-1995-y

Saxton, J., Morrow, L., Eschman, A., Archer, G., Luther, J., & Zuccolotto, A. (2009). Computer assessment of mild cognitive impairment. *Postgrad Med, 121*(2), 177-185. doi:10.3810/pgm.2009.03.1990

Scheltens, P., Blennow, K., Breteler, M. M., de Strooper, B., Frisoni, G. B., Salloway, S., & Van der Flier, W. M. (2016). Alzheimer's disease. *Lancet, 388*(10043), 505-517. doi:10.1016/S0140-6736(15)01124-1

Schmid, N. S., Ehrensperger, M. M., Berres, M., Beck, I. R., & Monsch, A. U. (2014). The extension of the German CERAD neuropsychological assessment battery with tests assessing subcortical, executive and frontal functions improves accuracy in dementia diagnosis. *Dement Geriatr Cogn Dis Extra, 4*(2), 322-334. doi:10.1159/000357774

Silbert, B., Evered, L., Scott, D. A., McMahon, S., Choong, P., Ames, D., . . . Jamrozik, K. (2015). Preexisting cognitive impairment is associated with postoperative cognitive dysfunction after hip joint replacement surgery. *Anesthesiology, 122*(6), 1224-1234. doi:10.1097/ALN.0000000000000671

Sprung, J., Roberts, R., Weingarten, T., Nunes Cavalcante, A., Knopman, D., Petersen, R., . . . Warner, D. (2017). Postoperative delirium in elderly patients is associated with subsequent cognitive impairment. *Br J Anaesth, 119*(2), 316-323.

Steinmetz, J., Christensen, K. B., Lund, T., Lohse, N., & Rasmussen, L. S. (2009). Long-term consequences of postoperative cognitive dysfunction. *Anesthesiology, 110*(3), 548-555. doi:10.1097/ALN.0b013e318195b569

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia, 47*(10), 2015-2028. doi:10.1016/j.neuropsychologia.2009.03.004

Story, D. A., Leslie, K., Myles, P. S., Fink, M., Poustie, S. J., Forbes, A., . . . New Zealand College of Anaesthetists Trials, G. (2010). Complications and mortality in older surgical patients in Australia and New Zealand (the REASON study): a multicentre, prospective, observational study. *Anaesthesia, 65*(10), 1022-1030. doi:10.1111/j.1365-2044.2010.06478.x

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests. Administration, Norms, and Commentary.* Oxford, UK.: Oxford University Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol., 18*, 643-662. doi:10.1037/h0054651

Summers, M. J., & Bondi, M. W. (2017). Editorial Introduction to the Special Issue on Neuropsychological Assessment in Aging. *Neuropsychol Rev, 27*(4), 303-304. doi:10.1007/s11065-017-9364-2

Thalmann, B., Spiegel, R., Stähelin, H. B., Brubacher, D., Ermini-Fünfschilling, D., Bläsi, S., & Monsch, A. U. (2002). Dementia screening in general practice: optimised scoring for the clock drawing test. *Brain Aging, 2*(2), 36-43.

Thomann, A. E., Goettel, N., Hessler, J. B., Berres, M., Jahn, T., & Monsch, A. U. (2018). *Differences in normative samples of the Montreal Cognitive Assessment: Call for guidelines*. Poster presented at the International Neuropsychological Society (INS) 2018 Mid-Year Meeting, Praque, Czech Republic.

Thomann, A. E., Goettel, N., Monsch, R. J., Berres, M., Jahn, T., Steiner, L. A., & Monsch, A. U. (2018). The Montreal Cognitive Assessment: Normative Data from a German-Speaking Cohort and Comparison with International Normative Samples. *J Alzheimers Dis, 64*(2), 643-655. doi:10.3233/JAD-180080

Trevethan, R. (2017). Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and Pitfalls in Research and Practice. *Front Public Health, 5*, 307. doi:10.3389/fpubh.2017.00307

Trzepacz, P. T., Hochstetler, H., Wang, S., Walker, B., Saykin, A. J., & Alzheimer's Disease Neuroimaging, I. (2015). Relationship between the Montreal Cognitive Assessment and Mini-mental State Examination for assessment of mild cognitive impairment in older adults. *BMC Geriatr, 15*, 107. doi:10.1186/s12877-015-0103-3

Trzepacz, P. T., Mittal, D., Torres, R., Kanary, K., Norton, J., & Jimerson, N. (2001). Validation of the Delirium Rating Scale-revised-98: comparison with the

delirium rating scale and the cognitive test for delirium. *J Neuropsychiatry Clin Neurosci, 13*(2), 229-242. doi:10.1176/jnp.13.2.229

van Meenen, L. C., van Meenen, D. M., de Rooij, S. E., & ter Riet, G. (2014). Risk prediction models for postoperative delirium: a systematic review and meta-analysis. *J Am Geriatr Soc, 62*(12), 2383-2390. doi:10.1111/jgs.13138

Vermeersch, P., Vercammen, M., Holvoet, A., Vande Broek, I., Delforge, M., & Bossuyt, X. (2009). Use of interval-specific likelihood ratios improves clinical interpretation of serum FLC results for the diagnosis of malignant plasma cell disorders. *Clin Chim Acta, 410*(1-2), 54-58. doi:10.1016/j.cca.2009.09.021

Weakley, A., Williams, J. A., Schmitter-Edgecombe, M., & Cook, D. J. (2015). Neuropsychological test selection for cognitive impairment classification: A machine learning approach. *J Clin Exp Neuropsychol, 37*(9), 899-916. doi:10.1080/13803395.2015.1067290

Weintraub, S., Besser, L., Dodge, H. H., Teylan, M., Ferris, S., Goldstein, F. C., . . . Morris, J. C. (2018). Version 3 of the Alzheimer Disease Centers' Neuropsychological Test Battery in the Uniform Data Set (UDS). *Alzheimer Dis Assoc Disord, 32*(1), 10-17. doi:10.1097/WAD.0000000000000223

Weir, A. J., Paterson, C. A., Tieges, Z., MacLullich, A. M., Parra-Rodriguez, M., Della Sala, S., & Logie, R. H. (2014). *Development of Android apps for cognitive assessment of dementia and delirium.* Paper presented at the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.

Weissberger, G. H., Strong, J. V., Stefanidis, K. B., Summers, M. J., Bondi, M. W., & Stricker, N. H. (2017). Diagnostic Accuracy of Memory Measures in Alzheimer's Dementia and Mild Cognitive Impairment: a Systematic Review and Meta-Analysis. *Neuropsychol Rev, 27*(4), 354-388. doi:10.1007/s11065-017-9360-6

Werner, F., Werner, K., & Oberzaucher, J. (2012). Tablets for seniors–an evaluation of a current model (iPad) *Ambient Assisted Living* (pp. 177-184). Heidelberg, Germany: Springer.

Wimo, A., Guerchet, M., Ali, G. C., Wu, Y. T., Prina, A. M., Winblad, B., . . . Prince, M. (2017). The worldwide costs of dementia 2015 and comparisons with 2010. *Alzheimers Dement, 13*(1), 1-7. doi:10.1016/j.jalz.2016.07.150

Winblad, B., Amouyel, P., Andrieu, S., Ballard, C., Brayne, C., Brodaty, H., . . . Zetterberg, H. (2016). Defeating Alzheimer's disease and other dementias: a priority for European science and society. *Lancet Neurol, 15*(5), 455-532. doi:10.1016/S1474-4422(16)00062-4

Witlox, J., Eurelings, L. S., de Jonghe, J. F., Kalisvaart, K. J., Eikelenboom, P., & van Gool, W. A. (2010). Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. *JAMA, 304*(4), 443-451. doi:10.1001/jama.2010.1013

Yaffe, K. (2018). Modifiable Risk Factors and Prevention of Dementia: What Is the Latest Evidence? *JAMA Intern Med, 178*(2), 281-282. doi:10.1001/jamainternmed.2017.7299

Yesavage, J. A., & Sheikh, J. I. (1986). Geriatric Depression Scale (GDS). *Clinical Gerontologist, 5*(1-2), 165-173. doi:10.1300/J018v05n01_09

Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer, 3*(1), 32-35.

Young, J., Meagher, D., & Maclullich, A. (2011). Cognitive assessment of older people. *BMJ, 343*, d5042. doi:10.1136/bmj.d5042

Zygouris, S., & Tsolaki, M. (2015). Computerized cognitive testing for older adults a review. *Am J Alzheimers Dis Other Dement, 30*(1), 13-28.

Curriculum Vitae

Alessandra Elena Thomann, born: December 1st, 1989 in Frauenfeld (TG)

Professional Experience

Since 01/2017	**PhD Student in Neuropsychology** Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel Department of Anesthesia, University Hospital Basel
01/2019 - 04/2019	**Clinical Neuropsychologist** Clienia Schlössli AG, Wetzikon
01/2016 - 12/2016	**Assistant Neuropsychologist** Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel
03/2015 - 12/2015 11/2013 - 08/2014	**Research Assistant** Memory Clinic, University Department of Geriatric Medicine FELIX PLATTER, Basel
11/2012 - 02/2013	**Internship Neuropsychological Research** Department of Neurophysiology, University Hospital Basel
08/2008 - 08/2009	**Internship Business Administration** World-Wide Business Centres AG, Basel

Education

	Master of Science in Psychology
02/2014 - 12/2015	University of Zurich, Switzerland Cognitive Psychology and Cognitive Neuroscience
09/2012 - 06/2013	University of Strasbourg, France Cognitive and Clinical Neuropsychology
	Bachelor of Science in Psychology
09/2011 - 06/2012	University of Paris 13, France Psychophysiology
09/2009 - 06/2011	University of Basel, Switzerland

Continuing education

2018 INS mid-year meeting, Prague, CZ; Big Data in Life Sciences, University Basel, CH; Python for Psychologists, University Basel, CH.
2017 The Munich Brain Course, Munich, DE; 32nd GNP Meeting, Konstanz, DE; 6th Symposium on Behavioral Neurology, Lucerne, CH; AAIC Conference, London, UK.
2016 5th Symposium on Behavioral Neurology, Lucerne, CH; INS mid-year meeting, London, UK.

Publications

- **Thomann AE**, Goettel N, Monsch RJ, Berres M, Jahn T, Steiner LA, Monsch AU.
  The Montreal Cognitive Assessment: Normative Data from a German-speaking Cohort and Comparison with International Normative Samples. J Alzheimers Dis. 2018;64(2):643-655.
- Monsch RJ, Burckhardt AC, Berres M, **Thomann AE**, Ehrensperger MM, Steiner LA, Goettel N. *Development of a Novel Self-administered Cognitive Assessment Tool and Normative Data for Older Adults. J* Neurosurg Anesthesiol. 2019;31(2):218-226.
- Steiner LA, Monsch R, **Thomann AE**, Monsch AU, Goettel N. *Transiente und permanente kognitive Defizite nach chirurgischen Operationen*. Ther Umsch. 2017;74(7):384-388.
- **Thomann AE**, Ehrensperger MM, Monsch AU. *BrainCheck und BrainCoach in der Hausarzt Praxis.* Hausarzt Praxis, 2017.

Conference Posters

- **Thomann AE,** Goettel N, Hessler JB, Berres M, Jahn T, Monsch AU. *Differences in normative samples of the Montreal Cognitive Assessment: Call for guidelines.* International Neuropsychological Society 2018 Mid-Year Meeting, Prague, CZ.
- **Thomann AE,** Goettel N, Berres M, Ehrensperger MM, Leyhe T, Monsch AU. *German Normative Data for the Montreal Cognitive Assessment.* Alzheimer's Association International Conference (AAIC) 2018, Chicago, USA. (Presenting author: Monsch AU).
- Goettel N, Monsch R, **Thomann AE**, Berres M, Steiner LA. *Preoperative cognitive assessment: development of a self-administered screening tool and normative data for an elderly population.* Euroanaesthesia 2018, Copenhagen, DK. (Presenting author: Goettel N).
- **Thomann AE**, Mistridis P, Ehrensperger MM, Monsch AU. *BrainCoach: a program of cognitive activation in primary care settings.* Alzheimer's Association International Conference (AAIC) 2017, London, UK.

Project work

- Validation studies in patients with mild and major neurocognitive disorder (Montreal Cognitive Assessment, CogCheck)
- Prevention of cognitive decline with a new program to motivate older healthy individuals to engage in cognitive leisure activities (www.braincoach-program.ch)
- Validation study on delirium-risk-prediction after cardiac surgery using a novel self-administered computerized cognitive assessment tool (CogCheck)
- Study on the association between changes in cerebral gray matter volume and postoperative cognitive dysfunction following sevoflurane anesthesia (POCD-MRI)

Presentations

11/2018*: Normative study on the MoCA and prevention of cognitive decline.* (Presentation in German); Master seminar at the University Basel

11/2018: *Early detection of cognitive decline.* (Presentation in German); Oberwil

11/2018: *The BrainCoach-Program: Cognitive activity for the possible prevention of cognitive decline.* (Presentation in German); Continuing education series «Selected chapters of the university center for medicine of Aging"; Felix Platter Hospital, Basel

08/2018: *The BrainCoach-Program. A new program for the possible prevention of cognitive decline.* (Presentation in French); Advisory Board Vifor Pharma, Lausanne

07/2018: *BrainCoach – a Program of cognitive activation in primary care settings.*

International Neuropsychological Society (INS) 2018 Mid-Year Meeting, Prague, CZ

11/2017: *Use it or lose it! Cognitive activity for the possible prevention of cognitive decline.* (Presentation in German); "Bärner Xundheitstag" for Pro Senectute / Alzheimer's Association, Berne

10/2017: *Subjective Cognitive Decline, SCD: Definition, relevance and diagnostics.* (Presentation in French); Advisory Board Vifor Pharma, Geneva

10/2017: *The BrainCoach-Program. A new program for the possible prevention of cognitive decline.* (Presentation in French); Advisory Board Vifor Pharma, Geneva

09/2017: *Differences in the Montreal Cognitive Assessment (MoCA) in two samples of cognitively healthy older individuals.* (Presentation in German); 32nd Gesellschaft für Neuropsychologie (GNP) Meeting, Konstanz